Semiparametric Methods for Contrasting Times Between Successive Events

by

Xu Shu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biostatistics)
in The University of Michigan
2015

Doctoral Committee:

Professor Douglas E. Schaubel, Chair
Research Assistant Professor Sehee Kim
Research Professor Ananda Sen
Assistant Professor Min Zhang

ACKNOWLEDGEMENTS

I would like to express my special thanks and appreciation to my advisor Professor Douglas E. Schaubel. Thank you for being such a wonderful advisor and great mentor for me. Without your constant instructions, inspirations, and encouragements, this thesis will not have been made possible. I have learned a great deal from you. Professionally, your broad knowledge and timely advice helped me grow tremendously. Personally, your enthusiasm and passionate attitude towards research inspired me all the time. I am truly fortunate to work with you on one of the most important achievements of my life.

I would also like to thank my committee members, Professor Sehee Kim, Professor Ananda Sen, and Professor Min Zhang. It is my great honor to have you on my committee. Thank you for your valuable suggestions and insightful comments at various stages of my research. You have made this thesis a significantly better one.

My special appreciation also goes to Professor Douglas E. Schaubel, Professor Pratima Sharma, and Professor John C. Magee. Thank you for your generous support in funding me through my doctoral study. Through working with you on diverse transplantation topics, I have developed many essential skills in application, which will definitely become a precious asset in my future career.

Last but not least, I would like to thank my family and friends, for their unconditional love and support. I am especially grateful to my father Zibin Shu, my mother Hua Yang, and my husband Di Yang. Thank you for always being there, for encouraging me to pursue my goals, and for making me a better person.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

LIST OF TABLES

TABLE

LIST OF APPENDICES

APPENDIX

ABSTRACT

Times between successive events (i.e., gap times) are of great importance in survival analysis. Although many methods exist for estimating covariate effects on gap times, very few existing methods allow for comparisons between gap times themselves. Motivated by the comparison of primary and repeat transplantation, our interest is specifically in contrasting the gap times. Two major challenges in gap time analysis are non-identifiability of the marginal distributions and the existence of dependent censoring (for all but the first gap time).

In the first chapter of this thesis, we propose methods to contrast gap time survival functions and their integration (restricted mean gap time). Specifically, we use Cox regression to estimate the (conditional) survival distributions of each gap time (given the previous gap times). Combining fitted survival functions based on those models, along with multiple imputation applied to censored gap times, we then contrast the first and second gap times with respect to average survival and restricted mean lifetime. Large-sample properties are derived, with simulation studies carried out to evaluate finite-sample performance. We apply the proposed methods to kidney transplant data obtained from a national organ transplant registry.

In the second chapter, we aim at contrasting gap time hazard functions, assuming that the hazard ratio between two hazard functions is constant over time. In particular, we propose a two-stage procedure, wherein the first stage involves a Cox regression model on the first gap time. Weighted estimating equations are then solved at the second stage to compare the first and second gap time hazard functions. We

derive the asymptotic properties of the proposed estimators, and investigate their performance in simulated finite samples. The proposed methods are applied to liver transplant data obtained from a national organ transplant registry.

The third chapter can be viewed as an extension of the second one, recognizing that the hazard ratio could be time-dependent. We propose semiparametric methods to estimate the gap time hazard functions, where the correlation between gap times are directly built into the model. Time-dependent hazard ratios can therefore be estimated, and the gap times are contrasted instantaneously. We also propose a novel form of average hazard ratio across time, leading to an overall contrast that is simple to interpret. Simulation studies are conducted under different scenarios to test the performance of our estimators in finite samples. The proposed methods are applied to liver transplant data obtained from a national registry to compare the hazard functions of post-transplant graft survival for first and second liver transplant.

Keywords: Gap times; Conditional model; Semiparametric model; Proportional hazards regression; Multiple imputation; Restricted mean lifetime; Weighted estimating equations; Hazard ratio; Time-dependent effect

CHAPTER I

Semiparametric Methods for Contrasting Gap Time Survival Functions

1.1 Introduction

In epidemiologic studies, a sequence of serial events is often of interest. Examples include numbered hospitalizations, tumor recurrences and, in a more general sense, transitions between states visited in a fixed order. There are two ways to define the time scale for multiple event data. The first, *total time*, measures time from a fixed time origin to an event. The second, *gap time*, measures time between successive events. However, for serial events, gap times are sometimes of more direct interest compared to total times, depending on the way the results are to be applied in practice. For instance, a patient who just got discharged from the hospital may question how long it will be until the next time he/she is hospitalized.

The analysis of gap times has several methodologic issues. Generally, the within-subject gap times are not independent. Even if total times are censored independently, the second and subsequent gap times will be subject to induced dependent censoring (Visser, 1996; Lin et al., 1999; Huang, 2000). For example, a longer first event time would normally indicate a larger probability of censoring for the second event. Thus, if the within-subject gap times are correlated, the second and subsequent gap times will depend on the censoring variables. This problem is one of two

major challenges in gap time analysis, with the second issue being non-identifiability. Specifically, when the support of the first gap time is not contained within the support of the censoring distribution, the marginal distributions of the second and subsequent gap times cannot be estimated nonparametrically, unless the gap times are independent (Lin et al., 1999; Wang, 1999; Huang, 2002; Schaubel and Cai, 2004a).

Gap time analysis has received much attention in the literature in recent years. Many nonparametric methods have been proposed, including Visser (1996), Wang and Wells (1998), Lin, Sun and Ying (1999), Wang and Chang (1999), Peña, Strawderman and Hollander (2001), van der Laan, Hubbard and Robins (2002), Schaubel and Cai (2004a), and Andrei and Murray (2006). The majority of these works developed nonparametric methods to estimate the joint and/or conditional distribution of the gap times. Semiparametric regression models have been proposed to account for covariate effects. The methods of Prentice, Williams and Peterson (1981) require that within-subject gap times are independent, conditional on the covariate. Huang (2002) proposed gap time regression methods based on the accelerated failure time model. Chen, Wang, and Huang (2004) proposed stratified proportional reverse-time hazards models to estimate a longitudinal patter of gap times. Schaubel and Cai (2004b) developed regression methods for the gap time hazard functions. Strawderman (2005) extended the accelerated failure time model for gap times that are independent conditional on the observed covariate. The method was subsequently extended to accommodate correlated gap times through a multiplicative gamma frailty (Strawderman, 2006). Huang and Liu (2007) used a joint frailty model to analyze disease recurrences and survival. Clement and Strawderman (2009) adapted generalized estimating equations to estimate the parameters indexing the conditional means and variances of gap times. Du, Jiang, and Wang (2011) proposed a smoothing spline

analysis of variance frailty model to estimate the gap time hazard.

Most existing regression methods target covariate effects within gap time, as opposed to contrasts between gap times themselves. One could append the first and second gap time data sets, then fit a marginal (common baseline) Cox (1972) model with an indicator for second gap time (first gap time then serving as the reference). This could be interpreted as a version of the Prentice et al. (PWP; 1981) method, or a form of the Wei, Lin and Weissfeld (1989) approach. Such a procedure (which is clearly not in line with the intended use of either PWP or WLW) would be biased due to failing to address either the previously mentioned identifiability or induced dependent censoring issues. One could not fit a frailty version of this model since there would be no repeated events within-individual (with covariate fixed; e.g., with an indicator for second-gap-time included in the covariate vector), and hence no information to estimate the frailty variance. The methods proposed by Chen, Wang, and Huang (2004) could be used to compare gap times, through an estimated longitudinal pattern parameter. Such longitudinal pattern could describe quantitatively the increasing or decreasing trend in gap times; however, a monotone trend is required. In addition, the subject-specific baseline hazard functions are sometimes unidentifiable. In summary, there are very few methods in the existing literature for comparing gap times, and the obvious extensions to existing methods have substantial limitations.

In this report, we develop methods to compare the survival functions and restricted mean lifetimes (Irwin, 1949) of the first (T_{i1}) and the second gap times ($\widetilde{T}_{i2}$). In particular, we contrast the average survival function for $\widetilde{T}_{i2}$ (obtained through appropriate conditioning, such as to respect the afore-described identifiability issues) with the corresponding survival function for T_{i1} obtained through the same averaging. The method we propose does not require inverse weighting or frailty modeling,

and works around the issues of non-identifiability and dependent censoring through flexible assumptions regarding the association between T_{i1} and $\widetilde{T}_{i2}$. Specifically, we contrast first and second gap time survival functions, as well as their integration over $[0, L]$ for pre-specified L. This difference in restricted mean survival times has been studied by many authors in various contexts (e.g., Karrison, 1987; Zucker, 1998; Chen and Tsiatis, 2001; Zhang and Schaubel, 2011; Zhang and Schaubel, 2012) and is a useful alternative to the hazard ratio.

Methods in this report are motivated by comparisons of graft survival between primary versus repeat kidney transplantation. This is a controversial research question of great interest to transplant surgeons and patients, which cannot be accurately addressed using existing gap time regression methods. The preferred therapy for patients with end-stage renal disease (ESRD) is kidney transplantation, due to increased survival and quality of life compared to the alternative, dialysis. Re-transplantation may be required if the original kidney transplant fails. In 2012, there were more than 104,000 patients on the waiting list for kidney transplantation, while the number of donor kidneys transplanted was approximately 13,000 (www.unos.org). Due to the relative scarcity of donor kidneys, it is meaningful to study whether patients with a repeat transplant have inferior outcomes compared to patients with a primary transplant. Such results would provide evidence to potentially serve as the basis for future organ allocation policy.

It has been frequently reported in the literature that graft survival is significantly lower for re-transplants relative to primary kidney transplants (Tejani and Sullivan, 1996; Pour-Reza-Gholi et al., 2005; Ahmed et al., 2008). However, there are also a few studies indicating that there was no significant difference (Gruber et al., 2009; Barba et al., 2011). Most of these articles used the Kaplan-Meier method (Kaplan and

Meier, 1958) and hence were not covariate-adjusted. In addition, the issues of non-identifiability and induced dependent censoring were not taken into consideration. Given the complexities of the data structure, much more detailed and robust analysis is required.

Our chief objective is to compare the average graft survival curve (i.e., an identifiable version thereof) for repeat kidney transplant patients, to the analogously averaged first-transplant survival curve. Rather than carry out predictions, our interest is in comparing primary and repeat transplant survival with respect to average survival, with the averaging being across which patients (i.e., as indexed by the covariate vector) receive a repeat transplant and when (in terms of follow-up time since first transplant). Consider the survival function for a re-transplant patient. Is this survival function really lower than that which would apply if in fact the patient were instead receiving a primary kidney transplant? If graft survival is truly lower for re-transplants, this would call into question the current policy of essentially assigning equal priority to primary and repeat kidney transplant candidates. Note that patient-specific contrasts between first and second transplant survival are at most of secondary interest; this is particularly true from a public health perspective, due to the impracticality of implementing patient-tailored organ allocation rules. In contrast, a global policy (applying to all patients) is feasible and is more likely to be perceived as fair by surgeons, patients and the public; a natural manner of arriving at such a policy is through a central measure such as the mean.

The remainder of this report is organized as follows. In Section 1.2, we introduce the required notation, then describe the measures proposed to compare gap times and their corresponding estimation procedures. In Section 1.3, the asymptotic properties of the proposed estimators are derived. A simulation study is described in Section 1.4.

In Section 1.5, we apply the proposed methods to kidney transplant data obtained from a national registry. Some remarks and discussion are given in Section 1.6.

1.2 Proposed Methods

In this section, we begin by formalizing the data structure and issues described in Section 1.1. We then describe the contrast of interest. This is followed by a description of the assumed models and proposed estimation procedures.

1.2.1 Notation and data structure

We first introduce the requisite notation. Let T_{ij} denote the jth total time ($j = 1, 2$) for subject i ($i = 1, \cdots, n$). To make our description more concrete, suppose that we are interested in comparing the first two gap times, T_{i1} and $\widetilde{T}_{i2} = T_{i2} - T_{i1}$. The case of comparing three or more gap times will be discussed in Section 1.6. The censoring time of the ith subject is denoted as C_i. Hence, T_{i1} is potentially censored by C_i and $\widetilde{T}_{i2}$ is potentially censored by $\widetilde{C}_{i2} = C_i - T_{i1}$. We let $\mathbf{Z}_i$ denote a vector of covariates for the ith subject, measured at baseline. Time-dependent covariates will be discussed in Section 1.6. We let $\tau_1 = \sup\{t : P(C_i > t) > 0\}$ denote the upper bound of the support of the first gap time's censoring distribution, and $\tau_2 = \sup\{t : P(\widetilde{C}_{i2} > t) > 0\}$. We define the counting process, $N_{i1}(t) = I(T_{i1} \leq t \wedge C_i)$. As well, it is convenient to define $N_{\bullet 1}(t) = \sum_{i=1}^{n} N_{i1}(t)$.

Since T_{i1} and $\widetilde{T}_{i2}$ are not likely to be independent in most biomedical examples, the two challenges described in Section 1.1, induced dependent censoring and non-identifiability, remain in the context of existing non- and semiparametric methods. In particular, $\widetilde{T}_{i2}$ is censored by $\widetilde{C}_{i2} = C_i - T_{i1}$. Therefore, even if C_i is independent of both T_{i1} and T_{i2}, $\widetilde{T}_{i2}$ will still be censored by a variate with which it is correlated. Further, although $P(T_{i1} > t)$ is identifiable nonparametrically on $[0, \tau_1]$, $P(\widetilde{T}_{i2} > t)$

is not identifiable without assumptions on the nature of the association between T_{i1} and $\widetilde{T}_{i2}$. However, conditional survival functions for $\widetilde{T}_{i2}$ can be identified. For example, $P(\widetilde{T}_{i2} > t | T_{i1} \leq t_1)$ is identifiable for a pre-specified and fixed time point, t_1; but, only on $t \in [0, \tau_1 - t_1]$. This notion has been used by previous authors (Lin et al., 1999; Schaubel and Cai, 2004a), but restrictions on the inference are clear. Although it leads to a useful description of $\widetilde{T}_{i2}$ nonparametrically, this construct is of limited value with respect to comparing gap times and was not motivated by such comparisons.

1.2.2 Contrasting first and second gap times

In the way of background, recall that our objective is to compare the first and second gap times. As implied in Section 1.2.1, the subjects are not homogenous, each being indexed by a covariate vector, $\mathbf{Z}_i$. Although, correspondingly, it is possible that the contrast between first and second gap times interacts with $\mathbf{Z}_i$, our interest is primarily in the average contrast. Naturally, to be meaningful, the average taken across the second gap times must be consistent with that taken across the first gap times, such that confounding eliminated by incorporating $\mathbf{Z}_i$ is not reintroduced by the averaging. Since survival probability tends to be easily understood by clinical investigators, we choose to contrast the gap times through differences in the survival function, and the integration thereof (restricted mean gap times).

To further elaborate on our perspective, consider again the motivating example. We could take an appropriately defined average graft survival function for repeat kidney transplants. A specific covariate distribution was used in deriving this average, and the same distribution would be used to average over the covariate-specific graft survival function for first transplants. The difference could then be taken (to compute the difference in graft survival probability) and integrated (to compute

difference in expected number of years lived out of the next 10).

We now formalize the concepts described above, starting with the second gap time, $\widetilde{T}_{i2}$. As described in Section 1.1, we are unable to estimate $P(\widetilde{T}_{i2} > t|\mathbf{Z}_i)$, but can estimate $S_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1) \equiv P(\widetilde{T}_{i2} > t|\mathbf{Z}_i, T_{i1}, T_{i1} \leq \tau_1)$ for $t \in [0, \tau_2]$. Taking the area under the curve, we can estimate $\mu_{i2}(L|\mathbf{Z}_i, T_{i1}; \tau_1) \equiv \int_0^L S_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1)dt$, with $L \leq \tau_2$. We let $q_{i1}(u, \mathbf{Z}_i)$ denote the joint density of $(T_{i1}, \mathbf{Z}_i)$; we then define $\pi_{i1}(u, \mathbf{Z}_i; \tau_1)$ to be the corresponding joint density of across the observable region pertaining to the first gap time,

$$\pi_{i1}(u, \mathbf{Z}_i; \tau_1) \;=\; \frac{q_{i1}(u, \mathbf{Z})}{\int_{\mathbf{Z}} \int_0^{\tau_1} q_{i1}(u, \mathbf{Z}_i)du \; d\mathbf{Z}_i},$$

where $\int_{\mathbf{Z}}$ represents an integral of dimension equal to that of $\mathbf{Z}_i$. Taking an average across $\{T_{i1}, \mathbf{Z}_i : \; T_{i1} \leq \tau_1\}$, we obtain

$$S_2(t; \tau_1) \;=\; \int_{\mathbf{Z}} \int_0^{\tau_1} S_{i2}(t|\mathbf{Z}_i, u; \tau_1)\pi_{i1}(u, \mathbf{Z}_i; \tau_1)du \; d\mathbf{Z}_i, \tag{1.1}$$

Note that $\int_{\mathbf{Z}} \int_0^{\tau_1} \pi_{i1}(u, \mathbf{Z}_i; \tau_1)du \; d\mathbf{Z}_i = 1$ is a valid joint distribution of $\{(T_{i1}, \mathbf{Z}_i) : T_{i1} \in (0, \tau_1]\}$. Having defined an appropriate survival function, we can then take $\mu_2(L; \tau_1) = \int_0^L S_2(t; \tau_1)dt$. We compute the average survival for the first gap time by taking an average analogous to (1.1), which implies using

$$S_1(t; \tau_1) \;=\; \int_{\mathbf{Z}} \int_0^{\tau_1} S_{i1}(t|\mathbf{Z}_i)\pi_{i1}(u, \mathbf{Z}_i; \tau_1)du \; d\mathbf{Z}_i. \tag{1.2}$$

Note that, as defined in (1.2), $S_1(t; \tau_1) \neq P(T_{i1} > t|T_{i1} \leq \tau_1)$, which would not yield an appropriate comparison. The survival function $S_1(t; \tau_1)$ was derived specifically as an appropriate comparator to $S_2(t; \tau_1)$ and its utility is mostly tied to that purpose. In the context of the kidney transplant example, $S_2(t; \tau_1)$ represents that appropriately averaged survival function for second transplant patients. The quantity $S_1(t; \tau_1)$ reflects survival after first transplant, averaged across the $\mathbf{Z}_i$ component

used in the calculation of $S_2(t;\tau_1)$. In comparing (1.1) and (1.2), the only difference is the use of $S_{i2}(t|\mathbf{Z}_i, T_{i1}, T_{i1} \leq \tau_1)$ in the former and $S_{i1}(t|\mathbf{Z}_i)$ in the latter. Finally, difference in the average survival curves is denoted $\delta(t;\tau_1) = S_2(t;\tau_1) - S_1(t;\tau_1)$, with the area between the survival curves given by $\Delta(L;\tau_1) = \int_0^L \delta(t;\tau_1)dt$.

1.2.3 Assumed models and proposed estimators

We seek to compare the first and second gap times in a manner which allows us to use all of the observed event times and does not require inverse weighting, without imposing unrealistic or unverifiable modeling assumptions. We therefore keep the modeling within a framework where model checking and validation are well-established. Along those lines, we assume that the first gap time follows a proportional hazards model (Cox, 1972),

$$\lambda_{i1}(t|\mathbf{Z}_i) = \lambda_{01}(t)\exp\{\boldsymbol{\beta}_1'\mathbf{Z}_i\}, \tag{1.3}$$

where $\lambda_{i1}(t|\mathbf{Z}_i) = \lim_{\delta\to 0}\delta^{-1}P(t \leq T_{i1} < t+\delta | T_{i1} \geq t, \mathbf{Z}_i)$. We chose a proportional hazards model because it is commonly used in censored data; it is flexible; and model checking procedures are widely available (e.g., Klein and Moeschberger, 2003). To address identifiability issues, we build a connection between the first and second gap time. We choose to work with the hazard function for the conditional variate, $\{\widetilde{T}_{i2}|\mathbf{Z}_i, T_{i1}, T_{i1} \leq \tau_1\}$,

$$\lambda_{i2}(t|\mathbf{Z}_i, T_{i1};\tau_1) = \lim_{\delta\to 0}\delta^{-1}P(t \leq \widetilde{T}_{i2} \ < t+\delta | \widetilde{T}_{i2} \geq t, \mathbf{Z}_i, T_{i1}, T_{i1} \leq \tau_1).$$

In particular, we assume that this quantity follows the following proportional hazards model,

$$\lambda_{i2}(t;\mathbf{Z}_i, T_{i1};\tau_1) = \lambda_{02}(t;\tau_1)\exp\{\boldsymbol{\beta}_2'\mathbf{Z}_i + \boldsymbol{\phi}_2'\boldsymbol{f}(T_{i1})\}, \tag{1.4}$$

where $\boldsymbol{f}(x)$ is a parametric possibly vector valued function of x. Model (1.3) and (1.4) together allow one to not only quantify the covariate effects on the hazards of first and second gap times, but also quantify the connection between the gap times. Moreover, the connection is parameterized in a very flexible way in model (1.4), because $\boldsymbol{f}$ can take a large number of possible forms (e.g., polynomial, spline, etc.). In order to decide what form $\boldsymbol{f}$ should take, one common strategy is to break continuous T_{i1} into a categorial variable through a set of functions. The model would then be fitted with the categorical version of T_{i1} in order to determine an appropriate functional form for $\boldsymbol{f}$. Denote $\boldsymbol{\theta}_2 = (\boldsymbol{\beta}_2', \boldsymbol{\phi}_2')'$ and set $\mathbf{W}_i = (\mathbf{Z}_i', \boldsymbol{f}(T_{i1})')'$.

The parameters $\boldsymbol{\beta}_1$ from model (1.3) and $\boldsymbol{\theta}_2$ from model (1.4) can be estimated through partial likelihood (Cox, 1975); while Breslow (1972) estimators are available for $\Lambda_{01}(t)$ and $\Lambda_{02}(t;\tau_1)$. After fitting models (1.3) and (1.4), the corresponding subject-specific survival functions can be estimated as follows

$$\widehat{S}_{i1}(t|\mathbf{Z}_i) = \exp\{-\widehat{\Lambda}_{i1}(t|\mathbf{Z}_i)\} \qquad \widehat{\Lambda}_{i1}(t|\mathbf{Z}_i) = \widehat{\Lambda}_{01}(t)\exp\{\widehat{\boldsymbol{\beta}}_1'\mathbf{Z}_i\}$$

$$\widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1};\tau_1) = \exp\{-\widehat{\Lambda}_{i2}(t|\mathbf{Z}_i, T_{i1};\tau_1)\} \qquad \widehat{\Lambda}_{i2}(t|\mathbf{Z}_i, T_{i1};\tau_1) = \widehat{\Lambda}_{02}(t;\tau_1)\exp\{\widehat{\boldsymbol{\theta}}_2'\mathbf{W}_i\}.$$

Ultimately, we will be averaging over $\{(\mathbf{Z}_i, T_{i1}) : T_{i1} \leq \tau_1\}$. The observed-data version of such averaging will depend on the censoring distribution, which of course is undesirable. As such, we multiply impute censored T_{i1} values. Specifically, a total of M imputations will be generated such that, in each imputation m, we set $T_{i1}^m = T_{i1}$ for subjects with $T_{i1} < C_i$; otherwise, we impute T_{i1}^m from the truncated distribution,

$$\widehat{P}(T_{i1} > t|C_i, T_{i1} > C_i, \mathbf{Z}_i) \equiv \frac{\widehat{S}_{i1}(t|\mathbf{Z}_i)I(t > C_i)}{\widehat{S}_{i1}(C_i|\mathbf{Z}_i)}. \tag{1.5}$$

It is natural to use this distribution because the only information we have is that the imputed T_{i1}^m will be larger than the censoring time C_i. Owing to its nonparametric component, the survival function estimator is not defined after τ_1. However, as will be

described shortly, if $T_{i1}^m > \tau_1$, then subject i does not contribute to the computation of the average survival curve for either the first or second gap time. Note that we used an 'improper' imputation method, referred to as Type-B imputation by Wang and Robins (1998) and Robins and Wang (2000), which means the estimated parameters $\widehat{\boldsymbol{\beta}}_1$ and $\widehat{\Lambda}_{01}(t)$ used in (1.5) are only estimated once and held as fixed in the imputation algorithm. This precludes the use of the familiar techniques for variance estimation in the presence of multiple imputation (e.g., Little and Rubin, 2002), as will be seen later.

After imputing T_{i1}^m when $T_{i1} > C_i$ and setting $T_{i1}^m = T_{i1}$ when $T_i < C_i$, we can compute

$$\widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) = \exp\{-\widehat{\Lambda}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1)\},$$

for the subset $\{i : T_{i1}^m \leq \tau_1\}$, where $\widehat{\Lambda}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) = \widehat{\Lambda}_{02}(t; \tau_1) \exp\{\widehat{\boldsymbol{\theta}}_2' \mathbf{W}_i^m\}$, with $\mathbf{W}_i^m = (\mathbf{Z}_i', \boldsymbol{f}(T_{i1}^m)')'$.

In evaluating $S_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1)$, a natural comparison is with $S_{i1}(t|\mathbf{Z}_i)$, which suggests the contrast, $S_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) - S_{i1}(t|\mathbf{Z}_i)$. For instance, in the context of kidney transplantation, both $\mathbf{Z}_i$ and T_{i1} are known at the time of the second transplant. In advising a patient about to undergo retransplantation (second kidney transplant), the survival distribution for the second gap time is naturally important; but, also the patient would likely be interested in how their graft failure the second time around would be (given what is known at the time of re-transplant: $\mathbf{Z}_i$, T_{i1}) compared to the risk they faced before the first transplant (given $\mathbf{Z}_i$). Note that, for the first and second gap time, the conditioning is on all information known at the respective gap time origins.

This gives rise to two useful contrasts, namely,

$$\widehat{\delta}_i^m(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) = \widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) - \widehat{S}_{i1}(t|\mathbf{Z}_i) \tag{1.6}$$

(1.7)

$$\widehat{\Delta}_i^m(L|\mathbf{Z}_i, T_{i1}; \tau_1) = \int_0^L \{\widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) - \widehat{S}_{i1}(t|\mathbf{Z}_i)\}dt = \int_0^L \widehat{\delta}_i^m(t|\mathbf{Z}_i, T_{i1}^m; \tau_1)dt.$$

The contrast in (1.6) represents the estimated distance between the first and second gap time survival functions, while (1.7) reflects the area between the survival functions over $[0, L]$.

Note that the contrasts are, now, specifically for each subject. Following (1.2) and (1.1), the average difference in gap time survival is given by

$$\widehat{\delta}^m(t; \tau_1) = \int_{\mathbf{Z}} \int_0^{\tau_1} \widehat{\pi}_{i1}^m(u, \mathbf{Z}_i; \tau_1)\widehat{\delta}_i^m(t|\mathbf{Z}_i, u; \tau_1)du\, d\mathbf{Z}_i, \tag{1.8}$$

where $\widehat{\pi}_{i1}^m(u, \mathbf{Z}_i; \tau_1) = N_{\bullet 1}^m(\tau_1)^{-1} \int_0^{\tau_1} dI\{T_{i1}^m \leq u\}$ with $N_{\bullet 1}^m(\tau_1) = \sum_{i=1}^n I\{T_{i1}^m \leq \tau_1\}$. The difference between the restricted mean lifetimes is then estimated by $\widehat{\Delta}^m(L; \tau_1) = \int_0^L \widehat{\delta}^m(t; \tau_1)dt$. The final estimates are averages of the M estimators obtained through multiple imputation:

$$\widehat{\delta}(t; \tau_1) = M^{-1} \sum_{m=1}^M \widehat{\delta}^m(t; \tau_1) \qquad \widehat{\Delta}(L; \tau_1) = \int_0^L \widehat{\delta}(t; \tau_1)dt.$$

1.3 Asymptotic Properties

We begin by establishing counting processes corresponding to the observed gap times. Recall (Section 1.2.1) that we defined $N_{i1}(t) = I(T_{i1} \leq t \wedge C_i)$ and $N_{\bullet 1}(t) = \sum_{i=1}^n N_{i1}(t)$ corresponding to T_{i1}. With respect to the imputed T_{i1}, we also defined $N_{i1}^m(t) = I(T_{i1}^m \leq t)$ and $N_{\bullet 1}^m(t) = \sum_{i=1}^n N_{i1}^m(t)$. For the second gap time, $\{\widetilde{T}_{i2}|T_{i1} \leq \tau_1 \wedge C_i\}$, we now define $\widetilde{N}_{i2}(t) = I(\widetilde{T}_{i2} \leq t \wedge \widetilde{C}_{i2}, T_{i1} \leq \tau_1 \wedge C_i)$. The at-risk processes are given by $Y_{i1}(t) = I(T_{i1} \wedge C_i \geq t)$ and $\widetilde{Y}_{i2}(t) = I(\widetilde{T}_{i2} \wedge \widetilde{C}_{i2} \geq t, T_{i1} \leq \tau_1 \wedge C_i)$, respectively. Then, the pertinent zero-mean processes are given by $M_{i1}(t) = N_{i1}(t) - \int_0^t \lambda_{i1}(u)Y_{i1}(u)du$, and $\widetilde{M}_{i2}(t) = \widetilde{N}_{i2}(t) - \int_0^t \lambda_{i2}(u; \tau_1)\widetilde{Y}_{i2}(u)du$.

Next, it is useful to define the following quantities:

$$s_1^{(d)}(t,\boldsymbol{\beta}_1) = E[Y_{i1}(t)\mathbf{Z}_i^{\otimes d}\exp\{\boldsymbol{\beta}_1'\mathbf{Z}_i\}]$$

$$S_1^{(d)}(t,\boldsymbol{\beta}_1) = n^{-1}\sum_{i=1}^{n} Y_{i1}(t)\mathbf{Z}_i^{\otimes d}\exp\{\boldsymbol{\beta}_1'\mathbf{Z}_i\}$$

$$s_2^{(d)}(t,\boldsymbol{\theta}_2) = E[\widetilde{Y}_{i2}(t)\mathbf{W}_i^{\otimes d}\exp\{\boldsymbol{\theta}_2'\mathbf{W}_i\}|T_{i1}\le\tau_1\wedge C_i]$$

$$S_2^{(d)}(t,\boldsymbol{\theta}_2) = N_{\bullet 1}(\tau_1)^{-1}\sum_{i=1}^{n} N_{i1}(\tau_1)\widetilde{Y}_{i2}(t)\mathbf{W}_i^{\otimes d}\exp\{\boldsymbol{\theta}_2'\mathbf{W}_i\}$$

for $d=0,1,2$, where, for a vector $\mathbf{a}$, $\mathbf{a}^{\otimes 0}=1$, $\mathbf{a}^{\otimes 1}=\mathbf{a}$, and $\mathbf{a}^{\otimes 2}=\mathbf{a}\mathbf{a}'$. Additionally we define the following two quantities:

$$\overline{\boldsymbol{z}}_1(t,\boldsymbol{\beta}_1) = \frac{\mathbf{s}_1^{(1)}(t,\boldsymbol{\beta}_1)}{s_1^{(0)}(t,\boldsymbol{\beta}_1)} \quad \overline{\boldsymbol{z}}_2(t,\boldsymbol{\theta}_2) = \frac{\mathbf{s}_2^{(1)}(t,\boldsymbol{\theta}_2)}{s_2^{(0)}(t,\boldsymbol{\theta}_2)}.$$

We assume the following regularity conditions for $i=1,\cdots,n$, $0\le s\le\tau_1$, $0\le u\le\tau_2$:

(a) $\{N_{i1}(\cdot),\widetilde{N}_{i2}(\cdot),Y_{i1}(\cdot),\widetilde{Y}_{i2}(\cdot),\mathbf{Z}_i\}$ are independent and identically distributed;

(b) $E[Y_{i1}(s)]>0$ and $E[\widetilde{Y}_{i2}(u)]>0$;

(c) elements of $\mathbf{Z}_i$ are bounded almost surely.

(d) $\Lambda_{01}(s)<\infty$ and $\Lambda_{02}(u;\tau_1)<\infty$

(e) positive-definiteness of the following matrices:

$$\begin{aligned}
\boldsymbol{\Sigma}_1(\boldsymbol{\beta}) &= E\left[\int_0^{\tau_1}\left\{\frac{\mathbf{s}_1^{(2)}(t,\boldsymbol{\beta}_1)}{s_1^{(0)}(t,\boldsymbol{\beta}_1)} - \overline{\boldsymbol{z}}_1(t,\boldsymbol{\beta}_1)^{\otimes 2}\right\}dN_{i1}(t)\right],\\
\boldsymbol{\Sigma}_2(\boldsymbol{\theta}) &= E\left[\int_0^{\tau_1}\pi_{i1}(u,\mathbf{Z}_i;\tau_1)du\int_0^{\tau_2}\left\{\frac{\mathbf{s}_2^{(2)}(t,\boldsymbol{\theta}_2)}{s_2^{(0)}(t,\boldsymbol{\theta}_2)} - \overline{\boldsymbol{z}}_2(t,\boldsymbol{\theta}_2)^{\otimes 2}\right\}d\widetilde{N}_{i2}(t)\right],
\end{aligned}$$

Before we go into the asymptotic properties of $n^{1/2}\{\widehat{\delta}(t;\tau_1)-\delta(t;\tau_1)\}$ and $n^{1/2}\{\widehat{\Delta}(L;\tau_1)-$

$\Delta(L;\tau_1)\}$, we lay out the following useful decompositions:

$$n^{1/2}[\widehat{\delta}(t;\tau_1)-\delta(t;\tau_1)] =$$

$$n^{1/2}M^{-1}\sum_{m=1}^{M} N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{i=1}^{n} N_{i1}^{m}(\tau_1)[\widehat{S}_{i2}(t|\mathbf{Z}_i,T_{i1}^{m};\tau_1)-S_2(t;\tau_1)] \tag{1.9}$$

$$-n^{1/2}M^{-1}\sum_{m=1}^{M} N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{i=1}^{n} N_{i1}^{m}(\tau_1)[\widehat{S}_{i1}(t|\mathbf{Z}_i)-S_1(t;\tau_1)]. \tag{1.10}$$

$$n^{1/2}[\widehat{\Delta}(L;\tau_1)-\Delta(L;\tau_1)] =$$

$$n^{1/2}M^{-1}\sum_{m=1}^{M} N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{i=1}^{n} N_{i1}^{m}(\tau_1)\int_0^L [\widehat{S}_{i2}(t|\mathbf{Z}_i,T_{i1}^{m};\tau_1)-S_2(t;\tau_1)]dt \tag{1.11}$$

$$-n^{1/2}M^{-1}\sum_{m=1}^{M} N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{i=1}^{n} N_{i1}^{m}(\tau_1)\int_0^L [\widehat{S}_{i1}(t|\mathbf{Z}_i)-S_1(t;\tau_1)]dt. \tag{1.12}$$

For convenience we define,

$$\begin{aligned}
(1.9) &= n^{1/2}[\widehat{S}_2(t;\tau_1)-S_2(t;\tau_1)]\\
(1.10) &= n^{1/2}[\widehat{S}_1(t;\tau_1)-S_1(t;\tau_1)]\\
(1.11) &= n^{1/2}[\widehat{\mu}_2(L;\tau_1)-\mu_2(L;\tau_1)]\\
(1.12) &= n^{1/2}[\widehat{\mu}_1(L;\tau_1)-\mu_1(L;\tau_1)],
\end{aligned}$$

where we define,

$$\begin{aligned}
\widehat{S}_2(t;\tau_1) &= M^{-1}\sum_{m=1}^{M} N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{i=1}^{n} N_{i1}^{m}(\tau_1)\widehat{S}_{i2}(t|\mathbf{Z}_i,T_{i1}^{m};\tau_1)\\
\widehat{S}_1(t;\tau_1) &= M^{-1}\sum_{m=1}^{M} N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{i=1}^{n} N_{i1}^{m}(\tau_1)\widehat{S}_{i1}(t|\mathbf{Z}_i)\\
\widehat{\mu}_2(L;\tau_1) &= \int_0^L \widehat{S}_2(t;\tau_1)dt\\
\widehat{\mu}_1(L;\tau_1) &= \int_0^L \widehat{S}_1(t;\tau_1)dt.
\end{aligned}$$

The asymptotic properties of (1.9), (1.10), (1.11), and (1.12) can be summarized by the following two theorems. Detailed proof is given in the Supplementary Materials in Appendix A.

THEOREM 1.1: *Under conditions (a) to (e), (1.9) and (1.11) have linear representations asymptotically; i.e.,*

$$n^{1/2}[\widehat{S}_2(t;\tau_1) - S_2(t;\tau_1)] = n^{-1/2}\sum_{i=1}^{n}\varphi_{i1}(t) + o_p(1);$$

$$n^{1/2}[\widehat{\mu}_2(L;\tau_1) - \mu_2(L;\tau_1)] = n^{-1/2}\sum_{i=1}^{n}\varphi_{i2}(L) + o_p(1),$$

where $\varphi_{i1}(t)$ and $\varphi_{i2}(L)$ $(i = 1, \cdots, n)$ are independent and identically distributed mean-zero random variables, such that $E\{\varphi_{i1}(t)^2\} < \infty$, $E\{\varphi_{i2}(L)^2\} < \infty$. Thus, (1.9) and (1.11) are asymptotically normal with mean 0 and variances $E\{\varphi_{i1}(t)^2\}$ and $E\{\varphi_{i2}(L)^2\}$, respectively. Specifically,

$$\begin{aligned}\varphi_{i1}(t) =& M^{-1}\sum_{m=1}^{M}\big\{P(T_{i1} \le \tau_1 \wedge C_i)^{-1}N_{i1}(\tau_1)\phi_{i1}^m(t) \\ &+ P(T_{i1} \le \tau_1)^{-1}N_{i1}^m(\tau_1)[S_{i2}(t|\mathbf{W}_i^m;\tau_1) - S_2(t;\tau_1)]\big\},\end{aligned}$$

$$\varphi_{i2}(L) = \int_0^L \varphi_{i1}(t)dt,$$

where we define

$$\begin{aligned}\phi_{i1}^m(t) &= \mathbf{b}_m(t)'\int_0^{\tau_2}[\mathbf{W}_i^m - \overline{\boldsymbol{z}}_2(t,\boldsymbol{\theta}_2)]d\widetilde{M}_{i2}(t) \\ &\quad - E[S_{i2}(t|\mathbf{W}_i;\tau_1)\exp(\boldsymbol{\theta}_2'\mathbf{W}_i)|T_{i1}^m \le \tau_1]\int_0^t s_2^{(0)}(u,\boldsymbol{\theta}_2)^{-1}d\widetilde{M}_{i2}(u),\end{aligned}$$

where $S_{i2}(t|\mathbf{W}_i|T_{i1} \le \tau_1) = S_{i2}(t|\mathbf{Z}_i, T_{i1};\tau_1)$ by definition (since $\mathbf{W}_i$ is a function of $\mathbf{Z}_i$ and T_{i1}) and with

$$\begin{aligned}\mathbf{b}_m(t) &= \boldsymbol{\Sigma}_2(\boldsymbol{\theta})^{-1}\int_0^t\{\overline{\boldsymbol{z}}_2(u,\boldsymbol{\theta}_2)E[S_{i2}(t|\mathbf{W}_i;\tau_1)\exp(\boldsymbol{\theta}_2'\mathbf{W}_i)|T_{i1}^m \le \tau_1] \\ &\quad - E[\mathbf{W}_i S_{i2}(t|\mathbf{W}_i;\tau_1)\exp(\boldsymbol{\theta}_2'\mathbf{W}_i)|T_{i1}^m \le \tau_1]\}d\Lambda_{02}(u;\tau_1),\end{aligned}$$

and $\boldsymbol{\Sigma}_2(\boldsymbol{\theta})$ is as defined in Condition (e).

Consistent estimators of the asymptotic variances of (1.9) and (1.11) are given by $n^{-2}\sum_{i=1}^{n}\widehat{\varphi}_{i1}(t)^2$ and $n^{-2}\sum_{i=1}^{n}\widehat{\varphi}_{i2}(L)^2$ respectively, where

$$\begin{aligned}\widehat{\varphi}_{i1}(t) =& M^{-1}\sum_{m=1}^{M}\Big\{\widehat{P}(T_{i1}\le\tau_1\wedge C_i)^{-1}N_{i1}(\tau_1)\widehat{\phi}_{i1}^{m}(t)\\ &+\widehat{P}(T_{i1}\le\tau_1)^{-1}N_{i1}^{m}(\tau_1)[\widehat{S}_{i2}(t|\mathbf{Z}_i,T_{i1}^{m};\tau_1)-\widehat{S}_2(t;\tau_1)]\Big\},\\ \widehat{\varphi}_{i2}(L) =& \int_0^L\widehat{\varphi}_{i1}(t)dt,\end{aligned}$$

and

$$\begin{aligned}\widehat{\phi}_{i1}^{m}(t) &= \widehat{\mathbf{b}}_m(t)'\int_0^{\tau_2}[\mathbf{W}_i^m-\overline{\boldsymbol{Z}}_2(t,\widehat{\boldsymbol{\theta}}_2)]d\widehat{M}_{i2}(t)\\ &\quad-N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{j=1}^{n}[N_{j1}^{m}(\tau_1)\widehat{S}_{j2}(t|\mathbf{W}_j^m;\tau_1)\exp(\widehat{\boldsymbol{\theta}}_2'\mathbf{W}_j^m)]\int_0^t S_2^{(0)}(u,\widehat{\boldsymbol{\theta}}_2)^{-1}d\widehat{M}_{i2}(u),\end{aligned}$$

where $d\widehat{M}_{i2}(t)=d\widetilde{N}_{i2}(t)-\widetilde{Y}_{i2}(t)d\widehat{\Lambda}_{i2}(t;\tau_1)$, and with

$$\begin{aligned}\widehat{\mathbf{b}}_m(t) &= \widehat{\boldsymbol{\Sigma}}_2(\boldsymbol{\theta})^{-1}\int_0^t\{\overline{\boldsymbol{Z}}_2(u,\widehat{\boldsymbol{\theta}}_2)N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{j=1}^{n}[N_{j1}^{m}(\tau_1)\widehat{S}_{j2}(t|\mathbf{W}_j^m;\tau_1)\exp(\widehat{\boldsymbol{\theta}}_2'\mathbf{W}_j^m)]\\ &\quad-N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{j=1}^{n}[N_{j1}^{m}(\tau_1)\mathbf{W}_j^m\widehat{S}_{j2}(t|\mathbf{W}_j^m;\tau_1)\exp(\widehat{\boldsymbol{\theta}}_2'\mathbf{W}_j^m)]\}d\widehat{\Lambda}_{02}(u;\tau_1),\end{aligned}$$

$$\widehat{\boldsymbol{\Sigma}}_2(\boldsymbol{\theta})=N_{\bullet 1}(\tau_1)^{-1}\sum_{i=1}^{n}N_{i1}(\tau_1)\int_0^{\tau_2}[\frac{\mathbf{S}_2^{(2)}(t,\widehat{\boldsymbol{\theta}}_2)}{S_2^{(0)}(t,\widehat{\boldsymbol{\theta}}_2)}-\overline{\boldsymbol{Z}}_2(t,\widehat{\boldsymbol{\theta}}_2)^{\otimes 2}]d\widetilde{N}_{i2}(t),$$

$$\overline{\boldsymbol{Z}}_2(t,\widehat{\boldsymbol{\theta}}_2)=\frac{\mathbf{S}_2^{(1)}(t,\widehat{\boldsymbol{\theta}}_2)}{S_2^{(0)}(t,\widehat{\boldsymbol{\theta}}_2)}.$$

THEOREM 1.2: *Under conditions (a) to (e), (1.10) and (1.12) have linear representations asymptotically; i.e.,*

$$n^{1/2}[\widehat{S}_1(t;\tau_1)-S_1(t;\tau_1)]=n^{-1/2}\sum_{i=1}^{n}\varphi_{i3}(t)+o_p(1);$$

$$n^{1/2}[\widehat{\mu}_1(L;\tau_1)-\mu_1(L;\tau_1)]=n^{-1/2}\sum_{i=1}^{n}\varphi_{i4}(L)+o_p(1),$$

where $\varphi_{i3}(t)$ and $\varphi_{i4}(L)$ $(i=1,\cdots,n)$ are independent and identically distributed mean-zero random variables, such that $E\{\varphi_{i3}(t)^2\}<\infty$, $E\{\varphi_{i4}(L)^2\}<\infty$. Thus,

(1.10) and (1.12) are asymptotically normal with means 0 and variances $E\{\varphi_{i3}(t)^2\}$ and $E\{\varphi_{i4}(L)^2\}$, respectively. Specifically,

$$\begin{aligned}\varphi_{i3}(t) =& M^{-1}\sum_{m=1}^{M}\left\{\mathbf{a}(t)'\int_0^{\tau_1}[\mathbf{Z}_i-\overline{\boldsymbol{z}}_1(t,\boldsymbol{\beta}_1)]dM_{i1}(t)-E[S_1(t|\mathbf{Z}_i)\exp(\boldsymbol{\beta}_1'\mathbf{Z}_i)]\right.\\ &\left.\times\int_0^t s_1^{(0)}(u,\boldsymbol{\beta}_1)^{-1}dM_{i1}(u)+P(T_{i1}\le\tau_1)^{-1}N_{i1}^m(\tau_1)[S_{i1}(t|\mathbf{Z}_i)-S_1(t;\tau_1)]\right\},\end{aligned}$$

$$\varphi_{i4}(L)=\int_0^L\varphi_{i3}(t)dt,$$

where

$$\mathbf{a}(t)=\boldsymbol{\Sigma}_1(\boldsymbol{\beta})^{-1}\int_0^t\{\overline{\boldsymbol{z}}_1(u,\boldsymbol{\beta}_1)E[S_1(t|\mathbf{Z}_i)\exp(\boldsymbol{\beta}_1'\mathbf{Z}_i)]-E[\mathbf{Z}_iS_1(t|\mathbf{Z}_i)\exp(\boldsymbol{\beta}_1'\mathbf{Z}_i)]\}d\Lambda_{01}(u),$$

and $\boldsymbol{\Sigma}_1(\boldsymbol{\beta})$ is as defined in Condition (e).

Consistent estimators of the asymptotic variances of (1.10) and (1.12) are $\sum_{i=1}^n\widehat{\varphi}_{i3}(t)^2/n^2$ and $\sum_{i=1}^n\widehat{\varphi}_{i4}(L)^2/n^2$ respectively, where

$$\begin{aligned}\widehat{\varphi}_{i3} =& M^{-1}\sum_{m=1}^{M}\left\{\widehat{\mathbf{a}}(t)'\int_0^{\tau_1}[\mathbf{Z}_i-\overline{\boldsymbol{Z}}_1(t,\widehat{\boldsymbol{\beta}}_1)]d\widehat{M}_{i1}(t)-n^{-1}\sum_{j=1}^n[\widehat{S}_{j1}(t|\mathbf{Z}_j)\exp(\widehat{\boldsymbol{\beta}}_1'\mathbf{Z}_j)]\right.\\ &\left.\times\int_0^t S_1^{(0)}(u,\widehat{\boldsymbol{\beta}}_1)^{-1}d\widehat{M}_{i1}(u)+\widehat{P}(T_{i1}\le\tau_1)^{-1}N_{i1}^m(\tau_1)[\widehat{S}_{i1}(t|\mathbf{Z}_i)-\widehat{S}_1(t;\tau_1)]\right\},\end{aligned}$$

with $d\widehat{M}_{i1}(t)=dN_{i1}(t)-Y_{i1}(t)d\widehat{\Lambda}_{i1}(t)$ and

$$\begin{aligned}\widehat{\mathbf{a}}(t) =& \widehat{\boldsymbol{\Sigma}}_1(\boldsymbol{\beta})^{-1}\int_0^t\left\{\overline{\boldsymbol{Z}}_1(u,\widehat{\boldsymbol{\beta}}_1)n^{-1}\sum_{j=1}^n[\widehat{S}_{j1}(t|\mathbf{Z}_j)\exp(\widehat{\boldsymbol{\beta}}_1'\mathbf{Z}_j)]\right.\\ &\left.-n^{-1}\sum_{j=1}^n[\mathbf{Z}_j\widehat{S}_{j1}(t|\mathbf{Z}_j)\exp(\widehat{\boldsymbol{\beta}}_1'\mathbf{Z}_j)]\right\}d\widehat{\Lambda}_{01}(u),\end{aligned}$$

$$\widehat{\boldsymbol{\Sigma}}_1(\boldsymbol{\beta})=n^{-1}\sum_{i=1}^n\int_0^{\tau_1}[\frac{\mathbf{S}_1^{(2)}(t,\widehat{\boldsymbol{\beta}}_1)}{S_1^{(0)}(t,\widehat{\boldsymbol{\beta}}_1)}-\overline{\boldsymbol{Z}}_1(t,\widehat{\boldsymbol{\beta}}_1)^{\otimes2}]dN_{i1}(t),$$

$$\overline{\boldsymbol{Z}}_1(t,\widehat{\boldsymbol{\beta}}_1)=\frac{\mathbf{S}_1^{(1)}(t,\widehat{\boldsymbol{\beta}}_1)}{S_1^{(0)}(t,\widehat{\boldsymbol{\beta}}_1)}.$$

The asymptotic linear representations of (1.9), (1.10), (1.11) and (1.12) follow from the large-sample results of Andersen and Gill (1982), assuming the imputation

model is correctly specified, and the fact that the limiting distribution of T_{i1}^m is the same as that of T_{i1}.

Combining the results in Theorem 1.1 and 1.2, the asymptotic properties for $n^{1/2}\{\widehat{\delta}(t;\tau_1) - \delta(t;\tau_1)\}$ and $n^{1/2}\{\widehat{\Delta}(L;\tau_1) - \Delta(L;\tau_1)\}$ can be readily summarized in the following theorem.

THEOREM 1.3: *Under conditions (a) to (c),* $n^{1/2}\{\widehat{\delta}(t;\tau_1)-\delta(t;\tau_1)\}$ *and* $n^{1/2}\{\widehat{\Delta}(L;\tau_1)-\Delta(L;\tau_1)\}$ *have linear representations asymptotically, i.e.,*

$$n^{1/2}\{\widehat{\delta}(t;\tau_1) - \delta(t;\tau_1)\} = n^{-1/2}\sum_{i=1}^{n}[\varphi_{i1}(t) - \varphi_{i3}(t)] + o_p(1);$$

$$n^{1/2}\{\widehat{\Delta}(L;\tau_1) - \Delta(L;\tau_1)\} = n^{-1/2}\sum_{i=1}^{n}[\varphi_{i2}(L) - \varphi_{i4}(L)] + o_p(1),$$

where $\varphi_{i1}(t), \varphi_{i2}(L), \varphi_{i3}(t), \varphi_{i4}(L), i = 1, \cdots, n$ are the same as above. Thus, $n^{1/2}\{\widehat{\delta}(t;\tau_1)-\delta(t;\tau_1)\}$ and $n^{1/2}\{\widehat{\Delta}(L;\tau_1) - \Delta(L;\tau_1)\}$are asymptotically normal with means 0 and variances $E\{[\varphi_{i1}(t) - \varphi_{i3}(t)]^2\}$, $E\{[\varphi_{i2}(L) - \varphi_{i4}(L)]^2\}$, respectively. The variances can be estimated as $n^{-2}\sum_{i=1}^{n}[\widehat{\varphi}_{i1}(t) - \widehat{\varphi}_{i3}(t)]^2$ and $n^{-2}\sum_{i=1}^{n}[\widehat{\varphi}_{i2}(L) - \widehat{\varphi}_{i4}(L)]^2$.

1.4 Simulations

We first describe the settings used in our simulation study. Each subject had two binary covariates Z_{i1} and Z_{i1} with $\Pr\{Z_{i1} = 1\}\Pr\{Z_{i2} = 1\} = 0.5$. For each subject, two gap times T_{i1} and $\widetilde{T}_{i2}$ were generated from the following proportional hazards models:

$$\lambda_{i1}(t) = \lambda_{01}(t)\exp\{\beta_1 Z_{i1} + \beta_2 Z_{i2}\}$$

$$\lambda_{i2}(t) = \lambda_{02}(t)\exp\{\beta_3 Z_{i1} + \beta_4 Z_{i2} + \beta_5 T_{i1}\}.$$

Parameter values used in Settings 1-4 are listed at the bottom of Table 1.1. The censoring time, C_i, followed a Uniform (0, 12) distribution. We set $L = \tau_1 = 5$, and

the number of multiple imputations to $M = 5$. The sample size was $n = 250$ for each data configuration, and we ran 1,000 replicates per configuration.

In Table 1.1, we present results from four parameter settings. In Settings 1-2, survival is much greater for T_{i1} then $\widetilde{T}_{i2}$, while the opposite is true for Settings 3-4. In each setting, bias is very small for both $\widehat{\Delta}(L;\tau_1)$ and the estimated survival probabilities. The estimated standard deviations (ESDs) and asymptotic standard errors (ASEs) match quite well, indicating that our asymptotic variance estimators are fairly accurate in reasonable size samples. Empirical coverage probabilities (ECP) are all around 0.95. The bias and discrepancy between the ASE and ESD are much larger for $\widehat{\Delta}(L;\tau_1)$ than those of the estimated survival functions, since restricted mean lifetimes can be viewed as an accumulation of survival probability, such that bias essentially propagates as t increases. Another thing to note is that the estimated survival probabilities at later time points are often more biased compared to those at earlier time points, which is intuitive because data are more sparse data towards the tail of the observation time distribution.

Additional data configurations are shown in the Supplementary Materials in Appendix A. Overall, the proposed methods are demonstrated to work well under the scenarios considered.

1.5 Application to kidney transplant data

We applied the proposed methods to kidney transplant data obtained from the Scientific Registry of Transplant Recipients (SRTR). The survival time of interest is time between kidney transplantation and graft failure, where graft failure is said to occur when the patient dies or the transplanted kidney ceases to function. A patient can have multiple kidney transplants if graft failure of the previous transplant(s)

Table 1.1: Simulation results for estimating survival functions and restricted mean lifetimes

	Setting 1					Setting 2				
Parameter	True	BIAS	ESD	ASE	ECP	True	BIAS	ESD	ASE	ECP
$\mu_1(L;\tau_1)$	3.08	-0.015	0.120	0.122	0.95	2.80	-0.007	0.123	0.128	0.95
$\mu_2(L;\tau_1)$	1.95	-0.007	0.154	0.146	0.94	1.77	-0.008	0.149	0.143	0.93
$\Delta(L;\tau_1)$	-1.13	0.008	0.189	0.195	0.95	-1.02	-0.001	0.175	0.199	0.97
$S_1(1;\tau_1)$	0.81	0.000	0.026	0.026	0.95	0.75	0.004	0.030	0.030	0.94
$S_2(1;\tau_1)$	0.62	0.004	0.044	0.043	0.94	0.56	0.001	0.044	0.043	0.95
$\delta(1;\tau_1)$	-0.19	0.004	0.050	0.051	0.94	-0.20	-0.004	0.051	0.051	0.95
$S_1(3;\tau_1)$	0.53	-0.001	0.034	0.034	0.95	0.46	-0.000	0.034	0.034	0.94
$S_2(3;\tau_1)$	0.25	0.003	0.044	0.041	0.93	0.22	0.004	0.039	0.039	0.94
$\delta(3;\tau_1)$	-0.28	0.004	0.054	0.054	0.95	-0.24	0.004	0.048	0.049	0.95
$S_1(5;\tau_1)$	0.35	0.001	0.034	0.034	0.95	0.30	0.001	0.030	0.030	0.96
$S_2(5;\tau_1)$	0.11	0.006	0.034	0.032	0.93	0.11	0.007	0.032	0.031	0.94
$\delta(5;\tau_1)$	-0.24	0.005	0.046	0.047	0.95	-0.19	0.006	0.041	0.041	0.95

	Setting 3					Setting 4				
Parameter	True	BIAS	ESD	ASE	ECP	True	BIAS	ESD	ASE	ECP
$\mu_1(L;\tau_1)$	2.13	-0.006	0.106	0.109	0.96	2.03	-0.008	0.101	0.111	0.96
$\mu_2(L;\tau_1)$	3.21	-0.008	0.146	0.141	0.94	3.01	-0.006	0.150	0.143	0.94
$\Delta(L;\tau_1)$	1.08	-0.002	0.168	0.178	0.96	0.98	0.002	0.162	0.185	0.97
$S_1(1;\tau_1)$	0.66	-0.001	0.030	0.031	0.95	0.61	-0.000	0.033	0.032	0.94
$S_2(1;\tau_1)$	0.82	-0.000	0.023	0.029	0.94	0.78	0.003	0.032	0.031	0.94
$\delta(1;\tau_1)$	0.16	0.001	0.040	0.041	0.96	0.17	0.003	0.032	0.031	0.94
$S_1(3;\tau_1)$	0.29	0.002	0.030	0.030	0.95	0.28	0.003	0.027	0.028	0.95
$S_2(3;\tau_1)$	0.56	0.003	0.039	0.034	0.95	0.51	0.002	0.041	0.040	0.94
$\delta(3;\tau_1)$	0.27	0.001	0.047	0.049	0.96	0.23	-0.001	0.046	0.046	0.94
$S_1(5;\tau_1)$	0.14	0.002	0.024	0.024	0.95	0.15	0.001	0.021	0.022	0.95
$S_2(5;\tau_1)$	0.39	0.004	0.044	0.043	0.94	0.36	0.003	0.043	0.041	0.94
$\delta(5;\tau_1)$	0.25	0.003	0.049	0.049	0.95	0.21	0.002	0.046	0.046	0.96

Setting 1: $\lambda_{01}(t)$=0.2, $\lambda_{02}(t)$=0.4, $\beta_1 = \beta_3 = \log(1.5)$, $\beta_2 = \beta_4 = -\log(1.5)$, $\beta_5 = \log(1.05)$
Setting 2: $\lambda_{01}(t)$=0.2, $\lambda_{02}(t)$=0.4, $\beta_1 = \beta_3 = \log(2.5)$, $\beta_2 = \beta_4 = -\log(2.5)$, $\beta_5 = \log(1.05)$
Setting 3: $\lambda_{01}(t)$=0.4, $\lambda_{02}(t)$=0.2, $\beta_1 = \beta_3 = \log(1.5)$, $\beta_2 = \beta_4 = -\log(1.5)$, $\beta_5 = -\log(1.05)$
Setting 4: $\lambda_{01}(t)$=0.4, $\lambda_{02}(t)$=0.2, $\beta_1 = \beta_3 = \log(2.5)$, $\beta_2 = \beta_4 = -\log(2.5)$, $\beta_5 = -\log(1.05)$

occurred. Our objective is to contrast the first and second transplants with respect to graft survival and restricted mean graft survival time. We included adult patients (age $\geq$18) who had their first transplant between January 1, 1998 and December 31, 2011. The observation period concluded on December 31, 2011. Kidney transplants are generally classified as deceased-donor or living-donor transplant. In our analysis we only include transplants from deceased donors.

Recipient-specific covariates include age at transplant, gender, race, diabetes status, body mass index (BMI), time waited for a transplant, calendar year of the transplant, and panel reactive antibodies (PRA). Covariates based on the the donor include donor age, BMI, serum creatinine, whether death was caused by stroke, hypertension, and diabetes status and duration. The covariate vector for each subject is recorded at each transplant and, hence, is transplant-specific. We estimated the functional form of $f(T_1)$ by first modeling a categorized version of T_1, in order to investigate the association pattern (discussed in Section 1.2). An appropriate function was determined to be $f(T_1) = (T_1 - 4)I(T_1 > 4)$, with time given in years.

Proportional hazards models were fitted for each of the two gap times. There are $n = 113,621$ subjects in total and the Cox model for the first gap time T_1 was fitted using $n_1 = 113,246$ subjects with no missing covariates at first transplant. Among the n_1 subjects, 39% are female; and 48% are white. The mean age at transplant is 52, with a standard deviation of 13. There are 39,817 graft failures or deaths after the first transplant.

We set $L = \tau_1 = 10$ years, which uses most of the available data while using values of τ_1 and L that would be meaningful to nephrologists and patients. There are 2,765 subjects with a second transplant that occurred within 10 years from the first transplant. Among those, $n_2 = 2,630$ subjects do not have any missing covariates at

second transplant and were used to fit a Cox model for the second gap time $\widetilde{T}_{i2}$. Of the n_2 subjects, 38% are female; and 52% are white. The mean age at transplant is 49, with a standard deviation of 13. There are 793 graft failures or deaths after the second transplant. As per the proposed methods, we average the primary- and repeat-transplant survival functions using the same covariate distribution. In cases where T_{i1}^m was imputed, the covariate from the primary transplant was used. Note that this would not affect the model fitting for re-transplanted patients, as it comes into play after model (1.4) has already been fitted. We used $M = 5$ multiple imputations.

Estimated average survival curves for first and second transplants are shown in Figure 1.1. The solid line is $\widehat{S}_2(t; 10)$; the dashed line is $\widehat{S}_1(t; 10)$, with t measured in years. The estimated survival function of the second gap time is below that of the first gap time for the first 8 years after transplant. However, the curves seem to get close and overlap after 8 years. The $\widehat{S}_1(t)$ curve is more smooth than $\widehat{S}_2(t)$, especially in the tail region, because the sample size is so much larger for the first gap time.

Estimated average 10-year mean graft survival times are contrasted in Table 1.2. Estimated restricted mean lifetime for the first gap time was $\widehat{\mu}_1(10; 10) = 6.916$ years with an estimated standard deviation of 0.014 years. The estimated restricted mean lifetime for the second gap time was $\widehat{\mu}_2(10; 10) = 6.588$ years with an estimated standard deviation of 0.144 years. The occurrence of this relatively large standard deviation is due to the respectively small sample size with respect to $\widetilde{T}_{i2}$. The difference between $\widehat{\mu}_2(10; 10)$ and $\widehat{\mu}_1(10; 10)$, $\widehat{\Delta}(10; 10)$, was -0.328 years (i.e., approximately 3.9 months), with an estimated standard deviation of 0.144 years. Thus, there is a statistically significant difference between first and second kidney transplants with respect to mean 10-year graft survival. However, the clinical importance

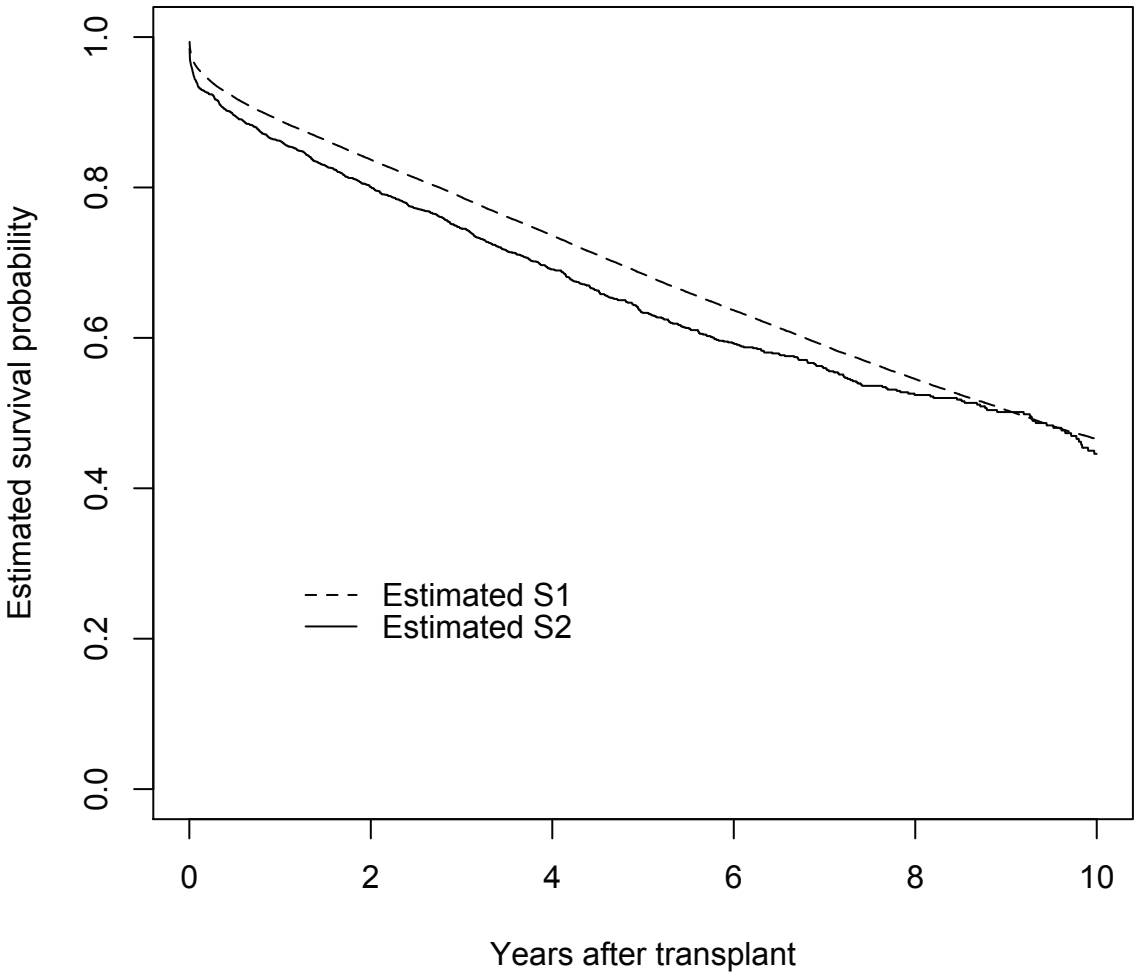

Figure 1.1: Analysis of SRTR data: Comparison of graft survival for first and second kidney transplants

Table 1.2: Analysis of SRTR data: Estimated 10-year mean graft survival time for first and second kidney transplants

Quantity	Estimate	Std. Error	p
$\mu_1(10;10)$	6.916	0.014	–
$\mu_2(10;10)$	6.588	0.144	–
$\Delta(10;10)$	-0.328	0.144	0.023

of a difference of 0.144 additional years is debatable.

We carried out various model diagnostics familiar to Cox regression (e.g., see Klein and Moeschberger, 2003). As an assessment of overall fit, the Cox-Snell residuals are plotted in Figure 1.2 for the first and second transplants (top left and top right panels, respectively). That the lines in each plot are approximately straight (except for the very end of follow-up) indicates no evidence of lack of fit in a general sense. Another concern was that the contrast in average graft survival would obscure important relationships at the patient-level. For example, it is possible that various covariates have important but opposite effects on the first and second transplant survival. This is assessed in Figure 1.2 (bottom left panel) where we plot scaled versions of the Z-scores for coefficients (scaled, to account for n_1 relative to n_2) from the T_{i1} and $\widetilde{T}_{i2}$ Cox models. Most points in this plot are in the upper right or lower left quadrant, indicating that the direction of the covariate effect is usually the same for the first and second transplant survival. Moreover, most points are close to the 45 degree line, indicating that the magnitude of the effect is typically quite similar for first and second transplants. A complete listing of parameter estimates and SEs for the first and second models is available from the Supplementary Materials in Appendix A. We also plotted a histogram (bottom right panel) of the patient-specific $\widehat{\Delta}_i(10;10)$ values use to compute the average effect (listed in Table 1.2). The distribution is bell-shaped with light tails and centered at $\approx$0. One would be concerned about using

Table 1.3: Naive analysis of SRTR data based on model: $\lambda_{i2}(t) = \lambda_0(t)\exp\{\boldsymbol{\beta}'\mathbf{Z}_i + \theta\}$

Gap time	$\widehat{\theta}$	SE($\widehat{\theta}$)	$\exp\{\widehat{\theta}\}$	p
1(ref.)	0	0	1	-
2	0.213	0.036	1.237	< 0.0001

the mean if it appeared (from such a histogram) to be heavily influenced by the tail. However, in our application, the mean does appear to represent the center of the data, with the majority of the $\widehat{\Delta}_i(10;10)$ values being with within $\pm$ 1 year. Hence, the mean is a reasonable summary measure in this application.

A naive way to contrast the first and second transplant would be to stack all data together, then fit a Cox regression with an indicator of re-transplant as a covariate. Doing so would be ignoring the induced dependent censoring and identifiability issues described in Section 1.1. Results from such an analysis of the model $\lambda_{i2}(t) = \lambda_0(t)\exp\{\boldsymbol{\beta}'\mathbf{Z}_i + \theta\}$ are summarized in Table 1.3. The estimated hazard ratio for re-transplant was 1.24 ($p < 0.0001$). As such, the post-second-transplant graft failure hazard would be interpreted as 24% higher than that of the post-first-transplant, which, in addition to being statistically significant, would be regarded as clinically important by most transplant surgeons and patients. The analysis in Table 1.3 is probably what would be carried out by the majority of data analysts unfamiliar with the statistical issues inherent to the gap time data structure.

1.6 Discussion

In this report, we propose semiparametric methods to compare the first and second gap times with respect to survival probability and restricted mean lifetime. Separate Cox models are assumed for the first and second gap times, with the first gap time used as a predictor of the hazard function for the second gap time. Multiple imputa-

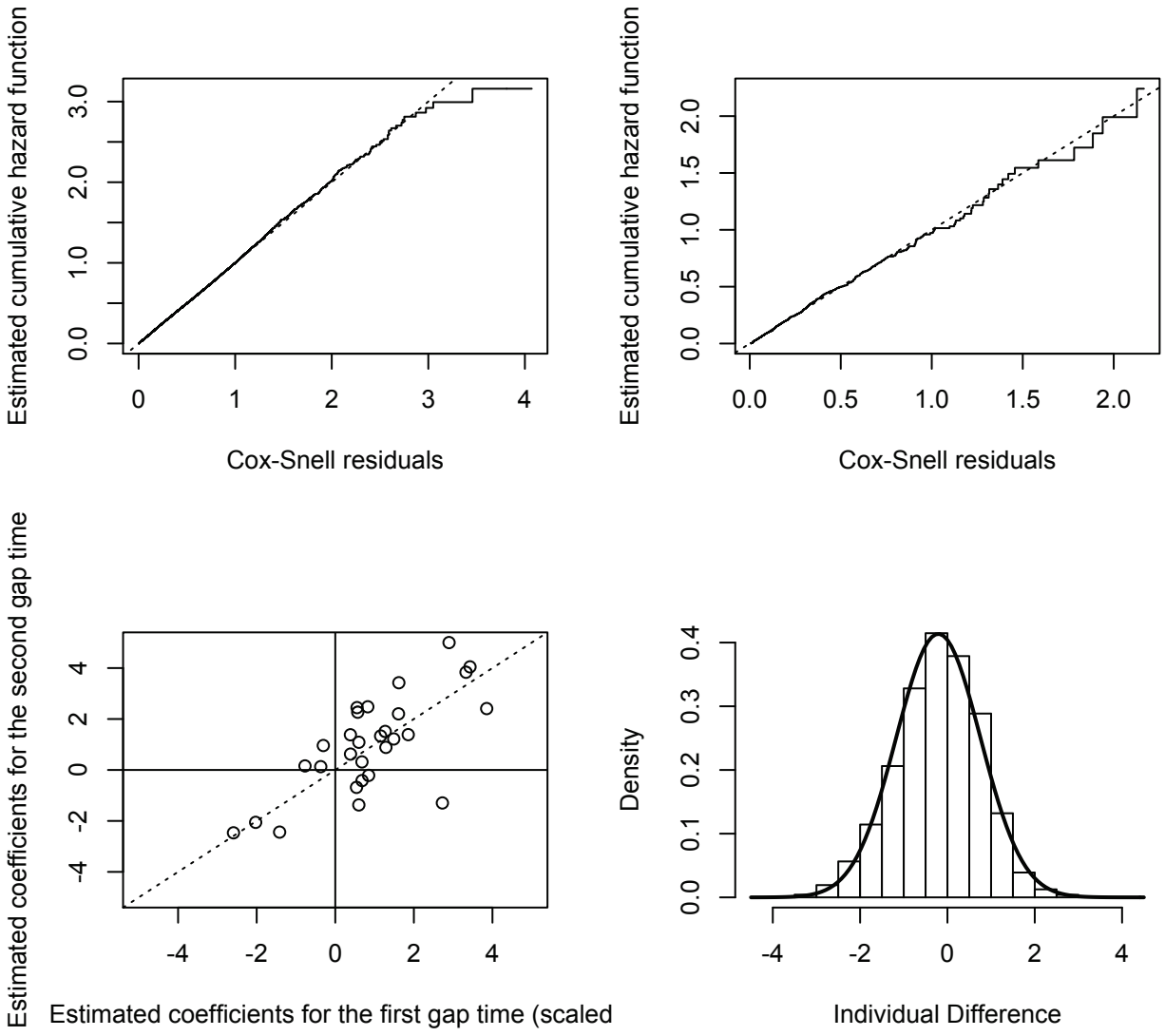

Figure 1.2: Diagnostic analysis of SRTR data: Cox-Snell residuals for first transplant model (top left) and second transplant model (top right); plot of coefficients for first vs. second transplant model (bottom left); histogram of patient-specific fitted difference in 10-year mean post-transplant survival time (bottom right)

tion of the first gap time is applied to identify subjects to be averaged in computing the survival probabilities and restricted mean lifetimes. Large-sample properties of the estimators are derived and demonstrated to work well in finite samples based on simulation studies.

We applied the proposed methods to compare the mean graft functioning lifetimes following first versus second kidney transplant, based on a 10-year time horizon. Our results imply that there is a significant difference between the two, which agrees with most existing studies (Tejani and Sullivan, 1996; Pour-Reza-Gholi et al., 2005; Ahmed et al., 2008) but contradicts with some recent analyses (Gruber et al., 2009; Barba et al., 2011). In contrast to results based on the proposed methods, results based on a model which simply used a second-transplant indicator show a significant 24% increase (Table 1.3) in the graft failure hazard associated with second kidney transplants, which would likely be viewed as clinically significant. Although statistically significant, the approximately 1/3 year difference in 10-year graft survival between primary and repeat kidney transplants is unlikely to be viewed as clinically important. The noteworthy difference in the results generated by the proposed method illustrates the importance of the method to the evaluation of re-transplantation.

The proposed methods entail to conditional inference on $\widetilde{T}_{i2}$ given $T_{i1} \leq \tau_1$. It is then required to pre-specify τ_1, which would typically be done before the analysis based on available follow-up and perhaps the general pattern of observed first and second gap time events. Under the proposed models and assumptions, one generally cannot identify the marginal distribution of $\widetilde{T}_{i2}$; the conditional distribution of the second gap time is identifiable for $T_{i1} \leq \tau_1$. Regarding the choice of τ_1, if one views marginal inference on $\widetilde{T}_{i2}$ as the gold standard, then the lower the choice of τ_1,

the more restricted is the resulting inference regarding the conditional distribution, $P(\widetilde{T}_{i2} > t|T_{i1} \leq \tau_1)$. In principle, τ_1 could be any value from zero to the largest observed censoring time. We would prefer τ_1 to be as large as possible so that we can include all available data into the analysis. So, the largest observed censoring time would be a choice that makes sense. In practice, it would often make more sense to choose a 'round' number that is very close to the largest observed censoring time, which can be readily grasped by the intended audience. For example in our case, we chose $\tau_1 = 10$ years.

The proposed methods are cast in terms of a baseline covariate, $\mathbf{Z}_i$. In settings where the covariates depend on follow up time, $\mathbf{Z}_i(t)$, one could model $\lambda_{i2}(t)$ as a function of $\{\mathbf{Z}_i(T_{i1}), T_{i1}\}$, which would be like a partly conditional model (Zheng and Heagerty, 2005; Gong and Schaubel, 2013). A complication would be that $\mathbf{Z}_i(T_{i1}^m)$ is not known in cases where $T_{i1} > C_i$, such that an imputed $\mathbf{Z}_i^m(T_{i1}^m)$ would be required. One possibility would be to apply longitudinal models of the time-dependent elements of $\mathbf{Z}_i(t)$.

For simplicity of exposition, in this report we focused on comparing the first two gap times. However, there can be situations where comparing three or more gap times is of interest; including the clinical setting that motivated our work, repeat kidney transplantation. One could generally follow the same framework we proposed.

Supplementary Materials

The proof of Theorems 1.1, 1.2, and 1.3, additional simulations that were carried out, and the parameter estimates for the gap-time-specific models are available in the Supplementary Materials in Appendix A, which can be found at the end of the three chapters.

CHAPTER II

Methods for Contrasting Gap Time Hazard Functions

2.1 Introduction

In clinical and biomedical studies, multiple event data are encountered often. Examples include the scenario where one person experiences a series of events, such as numbered hospitalizations. In this report, we are interested in the gap times; i.e., times between successive events. An alternative time scale for such data would be the time from the time origin to whenever the event occurs, known as the total time. However, gap times might be of more direct interest compared to total times, depending on the specific application. For example, a person who just received surgery to remove a cancerous tumor may wonder when s/he will experience a next tumor recurrence. Such questions are more readily addressed through a gap time scale.

Gap time analysis has been an active area of methodological research in recent years, partly because of its broad applicability to various fields. There are two major challenges that present difficulty in analyzing gap time data. The first challenge is non-identifiability (Lin et al., 1999; Wang, 1999; Huang, 2002; Schaubel and Cai, 2004a). In particular, because the support of the first gap time is usually not contained within the support of the censoring distribution, not all first gap times are

observed. As a result, the marginal distributions of the second and subsequent gap times cannot be identified without parametric assumptions, unless we assume that the within-subject gap times are independent of each other, an unreasonable assumption in most real-data applications. The second challenge is induced dependent censoring (Visser, 1996; Lin et al., 1999; Huang, 2000). For example, with a longer first gap time, the second gap time is more likely to be censored. Therefore, given that the gap times are typically correlated, the second and subsequent gap times will depend on the censoring variables, which violates the fundamental assumption of independent censoring.

Various methods have been proposed for gap time analysis. These existing methods can be broadly categorized into either those that estimate the joint/conditional survival functions non-parametrically, or those which incorporate the covariate effects on the gap time hazard functions semi-parametrically. The former category of methods includes those proposed by Visser (1996), Wang and Wells (1998), Lin, Sun and Ying (1999), Wang and Chang (1999), Peña, Strawderman and Hollander (2001), van der Lann, Hubbard and Robins (2002), Schaubel and Cai (2004a), and Andrei and Murray (2006). For the latter category, Prentice, Williams and Peterson (1981) assumed that the within-subject gap times are independent, and developed hazard models for each of the total and gap time scales. Huang (2002) developed regression methods for gap times based on the accelerated failure time model. Huang and Chen (2003) proposed a marginal proportional hazards model, when there are random effects leading to intra-subject correlation. Schaubel and Cai (2004b) proposed gap time regression methods for the hazard functions based on stratified proportional hazards models. Chen, Wang, and Huang (2004) proposed stratified proportional reverse-time hazard models, where a longitudinal pattern pa-

rameter can be estimated to compare gap times. However, a monotone time trend is required in the modeling; plus, the individual baseline hazard functions are sometimes not identifiable. Strawderman (2005) generalized the accelerated failure time model for gap times via conditional semiparametric intensity modeling. The model was later extended to handle correlated gap times by incorporating a multiplicative gamma frailty (Strawderman, 2006). Huang and Liu (2007) proposed a joint frailty model which could model disease recurrences and survival at the same time. Clement and Strawderman (2009) tried to estimate the parameters associated with the conditional means and variances of the gap times by modifying generalized estimating equations for longitudinal data. Du, Jiang, and Wang (2011) estimated the gap time hazard nonparametrically by building a smoothing spline ANOVA frailty model.

Almost none of these listed methods could be readily applied to compare gap times; those that could be used have rather strict assumptions leading to limited applicability. In this chapter, we propose novel methods for estimating the constant of proportionality assumed to connect the first and second gap times.

The motivating example for the proposed methods is in the context of liver transplantation. For patients with end-stage liver disease and acute liver failure, liver transplantation is the preferred treatment. Retransplantation may occur if the transplanted liver fails. However, there are many more (i.e., thousands more) patients on the wait-list compared to the number of donor organs available. Thus, it is extremely important to pursue an organ allocation policy that makes the best use of limited organs. Current policy uses the Model for End-Stage Liver Disease (MELD) score to determine a patient's ranking on the wait-list. The score is composed of serum creatinine, serum bilirubin, and international normalized ratio for prothrombin time (INR). It aims at predicting patients' survival on the waiting list. However, it does

not take into account whether this patient would be a primary (i.e., first-time) liver transplant recipient, or s/he has already had a liver transplant and, hence, is a repeat-transplant candidate. This motivates us to compare first and repeat (in particular, second) liver transplantation with respect to graft survival. Results of the analysis are crucial to the liver transplantation, from a public health perspective. For example, a finding that graft survival after the first transplant is significantly better than that after repeat transplantation might provide some new insights into how to assign the limited organs efficiently and fairly. It has been consistently reported that the outcomes of the repeat liver transplantation are inferior to those of the first transplantation (Markmann et al., 1997; Duran et al., 1998; Rosen & Martin, 1998; Ghobrial, 2002; Watt et al., 2003; Rao & Ojo, 2008; Kim et al., 2014). The most common method for estimating the graft survival after transplantation is Kaplan-Meier method. However, the issues of non-identifiability and dependent censoring discussed above were ignored, which may bias the comparison.

In this report, we propose methods for estimating the marginal hazard ratio connecting the first and second gap times. Specifically, a two-stage procedure is developed based on estimating equations. At the first stage, a proportional hazards model is fitted for the first gap time. Weighted estimating equations are then solved at the second stage to estimate the hazard ratio between the first and second gap times. The proposed estimator is intuitive and has a closed form. Rather than fitting a regression model to the second gap times, the second gap time counting process (a weighted version thereof) is used directly. The handling of the second gap time does not require additional assumptions regarding the between-gap-time association, but does in turn place restrictions on the range of the observed data that can be utilized (similar to various existing methods). For the case where the hazard functions of

gap times are no longer proportional to each other, we suggest reporting a sequence of truncated versions of proposed estimator, which would be evaluated at all jump times or perhaps a pre-specified grid of time points.

The methods we developed in Chapter I target at estimating and comparing the (conditional) survival functions and restricted mean lifetimes for gap times. Since hazard modeling is almost the default in survival analysis, here in this chapter, we propose methods to estimate the hazard ratio between gap time hazard functions. This serves as another perspective to contrast the gap times.

The remainder of the chapter is organized as follows. In Section 2.2, we describe the notation and methods. Asymptotic results are provided in Section 2.3. Simulation studies are carried out in Section 2.4 to demonstrate the performance of the estimator in finite samples. We apply the proposed methods to SRTR liver transplant data in Section 2.5. A discussion is given in Section 2.6.

2.2 Proposed Methods

First we introduce the necessary notation. Subject is denoted by $i = 1, \cdots, n$. For ease of presentation, we only consider the case of comparing two gap times. Suppose that $T_{ij} (j = 1, 2)$ are the total times of the events, such that T_{i1} is also the first gap time and $\widetilde{T}_{i2} = T_{i2} - T_{i1}$ is the second gap time. The censoring variable is C_i. So the first gap time is potentially censored by C_i, and the second gap time is potentially censored by $\widetilde{C}_{i2} = C_i - T_{i1}$. The covariate vector for subject i is given by $\mathbf{Z}_i$. We define $\tau = \sup\{t : P(C_i \geq t) > 0\}$. The counting and at-risk processes are defined

as,

$$
\begin{aligned}
N_{i1}(t) =& I(T_{i1} \wedge C_i \leq t, T_{i1} \leq C_i) \\
Y_{i1}(t) =& I(T_{i1} \wedge C_i \geq t) \\
N_{i2}(t; t_1) =& I(\widetilde{T}_{i2} \wedge \widetilde{C}_i \leq t, \widetilde{T}_{i2} \leq \widetilde{C}_i, T_{i1} \leq t_1) \\
Y_{i2}(t; t_1) =& I(\widetilde{T}_{i2} \wedge \widetilde{C}_i \geq t, T_{i1} \leq t_1).
\end{aligned} \tag{2.1}
$$

Because of the non-identifiability issues, we choose to work with the following hazard functions,

$$
\begin{aligned}
\lambda_{i1}(t) =& \lim_{\delta \to 0} \delta^{-1} P(t < T_{i1} \leq t + \delta | T_{i1} \geq t) \\
\lambda_{i2}(t; t_1) =& \lim_{\delta \to 0} \delta^{-1} P(t < \widetilde{T}_{i2} \leq t + \delta | \widetilde{T}_{i2} \geq t, T_{i1} \leq t_1).
\end{aligned}
$$

Then the conditional distribution of $\{\widetilde{T}_{i2} | T_{i1} \leq t_1\}$; including the hazard function $\lambda_{i2}(t; t_1)$, is identifiable for $t + t_1 \leq \tau$.

We assume that the gap time hazard functions are proportional, i.e.,

$$
\lambda_{i2}(t; t_1) = \lambda_{i1}(t) e^{\theta}.
$$

With this assumption, we set $t_1 = \tau/2$, so that both $\lambda_{i1}(t)$ and $\lambda_{i2}(t; t_1)$ can be compared for any $t \leq t_1$.

Referring back to the counting processes in (2.1), integrating and compensating produces the error terms,

$$
\begin{aligned}
M_{i1}(t) =& N_{i1}(t) - \int_0^t Y_{i1}(s) \Lambda_{i1}(ds) \\
M_{i2}(t; t_1) =& N_{i2}(t; t_1) - \int_0^t Y_{i2}(s; t_1) \Lambda_{i2}(ds; t_1),
\end{aligned}
$$

where $\Lambda_{i1}(t) = \int_0^t \lambda_{i1}(ds)$ and $\Lambda_{i2}(t; t_1) = \int_0^t \lambda_{i2}(ds; t_1)$.

In order to estimate θ, we propose to use a two-stage procedure. First, we model the hazard function of the first gap time with the following proportional hazards

model:

$$\lambda_{i1}(t) = \lambda_0(t) \exp\{\boldsymbol{\beta}' \mathbf{Z}_i\},$$

where $\lambda_0(t)$ is a baseline hazard function, and $\boldsymbol{\beta}$ is a parameter vector.

In the second stage, we solve the following weighted estimating equation,

$$\sum_{i=1}^{n} \int_0^t W_{i2}(s; t_1) M_{i2}(ds; t_1) = 0,$$

where we define the weight function $W_{i2}(s; t_1) = Y_{i2}(s; t_1) P(C_i \geq s + T_{i1} | T_{i1})^{-1}$. Solving the above equation, we have the following working estimator, applicable to any $t \leq t_1$,

$$\log \left\{ \frac{\sum_{i=1}^{n} \int_0^t W_{i2}(s; t_1) N_{i2}(ds; t_1)}{\sum_{i=1}^{n} \int_0^t W_{i2}(s; t_1) \Lambda_{i1}(ds; t_1)} \right\}. \tag{2.2}$$

We refer to this as a working estimator since it applies when both $\Lambda_{i1}(t)$ and $W_{i2}(s; t_1)$ are known, which is practically never the case. As implied previously, we compute $\widehat{\Lambda}_{i1}(t)$ using $\widehat{\boldsymbol{\beta}}$ and $\widehat{\Lambda}_0(t)$ from an unweighted Cox regression applied to the first gap time. It remains to estimate $P(C_i \geq s + T_{i1} | T_{i1})$, and we do so semi-parametrically.

To estimate the censoring distribution, we allow C_i to depend on a vector of covariates $\mathbf{Z}_i^C$ and, further, assume that the hazard of C_i follows the proportional hazards model,

$$\lambda_i^C(t) = \lambda_0^C(t) \exp\{\boldsymbol{\alpha}' \mathbf{Z}_i^C\}.$$

As such, we have $\widehat{P}(C_i > t | \mathbf{Z}_i^C) = \exp\{-\int_0^t \widehat{\lambda}_0^C(du) e^{\widehat{\boldsymbol{\alpha}}' \mathbf{Z}_i^C}\}$, while $\widehat{W}_{i2}(s; t_1) = Y_{i2}(s; t_1) \widehat{P}(C_i \geq s + T_{i1} | T_{i1})^{-1}$.

Thus, we can estimate θ by,

$$\widehat{\theta} = \log \left\{ \frac{\sum_{i=1}^{n} \int_0^{t_1} \widehat{W}_{i2}(s; t_1) N_{i2}(ds; t_1)}{\sum_{i=1}^{n} \int_0^{t_1} \widehat{W}_{i2}(s; t_1) \widehat{\Lambda}_{i1}(ds; t_1)} \right\}. \tag{2.3}$$

In the case of non-proportionality, one could contrast the first and second gap time hazards through $\widehat{\theta}(t)$, computed by replacing t_1 with t in the upper limit of the

integrals in (2.3). This exercise would be useful at least as a check of the essential assumption that the gap time hazards are proportional, and would indicate the nature of any violation in proportionality.

2.3 Asymptotic Properties

We define the following useful quantities:

$$
\begin{aligned}
s_C^{(d)}(t,\boldsymbol{\alpha}) =& E[Y_i^C(t)\boldsymbol{Z}_i^{C\otimes d}\exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}] \\
s_1^{(d)}(t,\boldsymbol{\beta}) =& E[Y_{i1}(t)\boldsymbol{Z}_i^{\otimes d}\exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}] \\
S_C^{(d)}(t,\boldsymbol{\alpha}) =& n^{-1}\sum_{i=1}^{n}[Y_i^C(t)\boldsymbol{Z}_i^{C\otimes d}\exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}] \\
S_1^{(d)}(t,\boldsymbol{\beta}) =& n^{-1}\sum_{i=1}^{n}[Y_{i1}(t)\boldsymbol{Z}_i^{\otimes d}\exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}] \\
\overline{\boldsymbol{z}}_C(t,\boldsymbol{\alpha}) =& \frac{\boldsymbol{s}_C^{(1)}(t,\boldsymbol{\alpha})}{s_C^{(0)}(t,\boldsymbol{\alpha})} \\
\overline{\boldsymbol{z}}_1(t,\boldsymbol{\alpha}) =& \frac{\boldsymbol{s}_1^{(1)}(t,\boldsymbol{\alpha})}{s_1^{(0)}(t,\boldsymbol{\alpha})}
\end{aligned}
$$

for $d = 0, 1, 2$, where, for a vector $\boldsymbol{a}$, $\boldsymbol{a}^{\otimes 0} = 1$, $\boldsymbol{a}^{\otimes 1} = \boldsymbol{a}$, and $\boldsymbol{a}^{\otimes 2} = \boldsymbol{a}\boldsymbol{a}'$.

The counting and at-risk processes of the censoring variable C_i are defined as,

$$
\begin{aligned}
N_i^C(t) =& I(C_i \le T_{i1} + \widetilde{T}_{i2}, C_i \le t) \\
Y_i^C(t) =& I(C_i \wedge (T_{i1} + \widetilde{T}_{i2}) \ge t).
\end{aligned}
$$

Thus, the error term is,

$$
M_i^C(t) = N_i^C(t) - \int_0^t Y_i^C(t)\Lambda_i^C(ds),
$$

where $\Lambda_i^C(t) = \int_0^t \lambda_i^C(ds)$.

In order to derive the asymptotic properties for $\widehat{\theta}$, we assume that the following regularity conditions hold for $i = 1, \cdots, n$, $0 \le s \le t_1$, $0 \le u \le t_1$:

(a) $\{N_{i1}(\cdot), N_{i2}(\cdot), Y_{i1}(\cdot), Y_{i2}(\cdot), \mathbf{Z}_i\}$ are independent and identically distributed;

(b) $E[Y_{i1}(s)] > 0$ and $E[Y_{i2}(u; t_1)] > 0$;

(c) elements of $\mathbf{Z}_i$ are bounded almost surely;

(d) $\Lambda_{01}(s) < \infty$ and $\Lambda_{02}(u; t_1) < \infty$;

(e) positive-definiteness of the following matrices:

$$\begin{aligned}\boldsymbol{\Sigma}_C(\boldsymbol{\alpha}) =& E\left[\int_0^\tau \left\{\frac{\boldsymbol{s}_C^{(2)}(t,\boldsymbol{\alpha})}{s_C^{(0)}(t,\boldsymbol{\alpha})} - \overline{\boldsymbol{z}}_C(t,\boldsymbol{\alpha})^{\otimes 2}\right\} dN_i^C(t)\right],\\ \boldsymbol{\Sigma}_1(\boldsymbol{\beta}) =& E\left[\int_0^\tau \left\{\frac{\boldsymbol{s}_1^{(2)}(t,\boldsymbol{\beta})}{s_1^{(0)}(t,\boldsymbol{\beta})} - \overline{\boldsymbol{z}}_1(t,\boldsymbol{\beta})^{\otimes 2}\right\} dN_{i1}(t)\right].\end{aligned}$$

The asymptotic properties is summarized by the following theorem. A detailed proof is given in the Supplementary Materials in Appendix B.

THEOREM 2.1: *Under conditions (a) to (e), we have the following linear representations pertinent to the proposed estimator,*

$$n^{1/2}(\widehat{\theta} - \theta) = n^{-1/2}\sum_{i=1}^{n}\phi_i(t) + o_p(1),$$

where $\phi_i(t)$ $(i = 1, \cdots, n)$ are independent and identically distributed mean-zero random variables, such that $E\{\phi_i(t)^2\} < \infty$. Thus, $n^{1/2}(\widehat{\theta} - \theta)$ is asymptotically normal with mean 0 and variance $E\{\phi_i(t)^2\}$, with

$$\phi_i(t) = e^{-\theta}[\frac{1}{B(t)}\varphi_{i1}(t) - \frac{A(t)}{B(t)^2}\varphi_{i2}(t)],$$

where

$$
\begin{aligned}
A(t) &= \int_0^t E[W_{i2}(s;t_1)\exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}]d\Lambda_{02}(s;t_1) \\
B(t) &= \int_0^t E[W_{i2}(s;t_1)\exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}]d\Lambda_{01}(s) \\
\varphi_{i1}(t) &= \int_0^t n^{-1}\sum_{j=1}^n W_{j2}(s)\left\{\boldsymbol{k}'_{jC}(s+T_{j1};\boldsymbol{\alpha})\boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1}\mathbf{U}_i^C(\boldsymbol{\alpha})\right. \\
&\quad \left.+\exp\{\boldsymbol{\alpha}'\mathbf{Z}_j^C\}\int_0^{s+T_{j1}} s_C^{(0)}(u;\boldsymbol{\alpha})^{-1}dM_i^C(u;\boldsymbol{\alpha})\right\}N_{j2}(ds;t_1) + \int_0^t W_{i2}(s;t_1)M_{i2}(ds;t_1) \\
\varphi_{i2}(t) &= \int_0^t n^{-1}\sum_{j=1}^n W_{j2}(s)\left\{\boldsymbol{k}'_{jC}(s+T_{j1};\boldsymbol{\alpha})\boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1}\mathbf{U}_i^C(\boldsymbol{\alpha})\right. \\
&\quad \left.+\exp\{\boldsymbol{\alpha}'\mathbf{Z}_j^C\}\int_0^{s+T_{j1}} s_C^{(0)}(u;\boldsymbol{\alpha})^{-1}dM_i^C(u;\boldsymbol{\alpha})\right\}\Lambda_{j1}(ds) \\
&\quad + \int_0^t \frac{1}{n}\sum_{j=1}^n W_{j2}(s;t_1)\exp\{\boldsymbol{\beta}'\mathbf{Z}_j\}\left\{(\mathbf{Z}_j-\overline{\boldsymbol{z}}_1(s;\boldsymbol{\beta}))'\boldsymbol{\Sigma}_1(\boldsymbol{\beta})^{-1}\mathbf{U}_{i1}(\boldsymbol{\beta})\Lambda_{01}(ds)\right. \\
&\quad \left.+s_1^{(0)}(s;\boldsymbol{\beta})^{-1}dM_{i1}(s;\boldsymbol{\beta})\right\},
\end{aligned}
$$

and

$$
\begin{aligned}
\boldsymbol{k}_{jC}(t;\boldsymbol{\alpha}) &= \exp\{\boldsymbol{\alpha}'\mathbf{Z}_j^C\}\int_0^t \left\{\mathbf{Z}_j^C - \overline{\mathbf{z}}_C(t;\boldsymbol{\alpha})\right\}d\Lambda_0^C(t) \\
\mathbf{U}_i^C(\boldsymbol{\alpha}) &= \int_0^\tau \{\boldsymbol{Z}_i^C - \overline{\boldsymbol{z}}_C(t;\boldsymbol{\alpha})\}dM_i^C(t;\boldsymbol{\alpha}) \\
\mathbf{U}_{i1}(\boldsymbol{\beta}) &= \int_0^\tau \{\boldsymbol{Z}_i - \overline{\boldsymbol{z}}_1(t;\boldsymbol{\beta})\}dM_{i1}(t;\boldsymbol{\beta}).
\end{aligned}
$$

The asymptotic linear representation follows from the large-sample results of Andersen and Gill (1982). Consistent estimator of the asymptotic variance is given by $n^{-2}\sum_{i=1}^n \widehat{\phi}_i(t)^2$, which is written out in the Supplementary Materials in Appendix B.

2.4 Simulations

We simulate $n = 250$ subjects, each with two covariates, Z_{i1} and Z_{i2}, which are independent Bernoulli(0.5) variables. Our objective is to simulate the first and

second gap times from the following two proportional hazards models:

$$P(T_{i1} > t) = \exp\{-\lambda_0 t e^{\boldsymbol{\beta}'\mathbf{Z}_i}\}$$

$$P(\widetilde{T}_{i2} > t | T_{i1} \leq t_1) = \exp\{-\lambda_0 t e^{\boldsymbol{\beta}'\mathbf{Z}_i + \theta}\},$$

and T_{i1} is correlated with $\widetilde{T}_{i2}$ given $\mathbf{Z}_i = \{Z_{i1}, Z_{i2}\}$.

The core idea we used to simulate two correlated variables is to generate two correlated uniformly distributed random variables first, and then use them to inverse the wanted survival functions. In order to simulate two correlated uniformly distributed random variables, we generate variables following a bivariate normal distribution with a correlation coefficient ρ, so that the marginal survival functions would both be $U(0,1)$. This idea is twisted somewhat because here we try to generate a marginal distribution (for T_{i1}) and a conditional distribution (for $\widetilde{T}_{i2}$).

Specifically, we first generate $\Delta_i = I\{T_{i1} \leq t_1\}$ for $i = 1, \cdots, n$ from a Bernoulli distribution with probability

$$P(\Delta_i = 1) = P(T_{i1} \leq t_1) = 1 - \exp\{-\lambda_0 t_1 e^{\boldsymbol{\beta}'\mathbf{Z}_i}\}.$$

For subjects with $\Delta_i = 1$, we simulate the first and second gap times from:

$$P(T_{i1} > t | T_{i1} \leq t_1) = \frac{\exp\{-\lambda_0 t e^{\boldsymbol{\beta}'\mathbf{Z}_i}\} - \exp\{\lambda_0 t_1 e^{\boldsymbol{\beta}'\mathbf{Z}_i}\}}{1 - \exp\{-\lambda_0 t_1 e^{\boldsymbol{\beta}'\mathbf{Z}_i}\}} \tag{2.4}$$

$$P(\widetilde{T}_{i2} > t | T_{i1} \leq t_1) = \exp\{-\lambda_0 t e^{\boldsymbol{\beta}'\mathbf{Z}_i + \theta}\} \tag{2.5}$$

For subjects with $\Delta_i = 0$, we simulate the first and second gap times from:

$$P(T_{i1} > t | T_{i1} > t_1) = \frac{\exp\{-\lambda_0 t e^{\boldsymbol{\beta}'\mathbf{Z}_i}\}}{\exp\{-\lambda_0 t_1 e^{\boldsymbol{\beta}'\mathbf{Z}_i}\}} \tag{2.6}$$

$$P(\widetilde{T}_{i2} > t | T_{i1} > t_1) = \exp\{-\lambda_0 t e^{\boldsymbol{\beta}'\mathbf{Z}_i}\} \tag{2.7}$$

To generate correlated T_{i1} and $\{\widetilde{T}_{i2} | T_{i1} \leq t_1\}$, we use the core idea discussed previously. Specifically, we first generate N_1 and N_2 from a bivariate normal distribution

with means 0, variances 1, and a correlation coefficient ρ. Then $U_1 = P(N_1 > t)$ and $U_2 = P(N_2 > t)$ are correlated random variables, and they both follow a uniform distribution on $(0, 1)$. Finally U_1 and U_2 are used to inverse the equations (2.4) and (2.5) to generate T_{i1} and T_{i2}. In equations (2.6) and (2.7), it is not necessary that we use this method because T_{i1} and $\{\widetilde{T}_{i2}|T_{i1} > t_1\}$ need not to be correlated. The hazard function of the censoring variable, C_i, follows a proportional hazards model:

$$P(C_i > t|\mathbf{Z}_i) = \exp\{-\lambda_0^C t e^{\boldsymbol{\alpha}'\mathbf{Z}_i^C}\}.$$

We set the parameters $\beta_1 = \log(1.5)$, $\beta_2 = -\log(1.5)$, and $\theta = 0.5$. Baseline hazard λ_0 is set to be 0.6. We use $t_1 = 3$ and $\rho = 0.5$. For the censoring parameters, $\lambda_0^C = 0.2$, $\alpha_1 = 0.2$, and $\alpha_2 = -0.2$. Under this data configuration, the probability of $\{T_{i1} \leq t_1\}$ is approximately 0.83. The correlation between T_{i1} and $\widetilde{T}_{i2}$ conditioning on $\{T_{i1} \leq t_1\}$ is 0.47. There are around 25% first gap times being censored, and, in subjects with $\{T_{i1} \leq t_1\}$, approximately 30% of the second gap times are censored.

We ran 1,000 replicates for this data configuration, with the estimated θ evaluated at three time points: $t = 1$, 2 and 3. The second row of Table 2.2 (Setting B) lists the results from the simulation. The biases are very close to zero for each time point we evaluated. The empirical standard deviations (ESDs) and the averaged asymptotic standard errors (ASEs) match very well across all time points. The estimated 95% coverage probability is around 0.95. The results indicate that the estimator works well under the scenario we considered. In addition, the asymptotic variance estimator is very accurate in estimating the standard errors in the reasonable-sized sample.

Additional scenarios are also simulated. Firstly, two extra cases with different correlation coefficients are considered. Specifically, one is with weaker correlation $\rho = 0.3$ (Setting A), and the other is with stronger correlation $\rho = 0.7$ (Setting C). Then, heavier censoring is simulated by re-setting $\lambda_0^C = 0.3$, which results in approximately

Table 2.1: Data configurations for simulations

Setting	λ_0^C	α_1	α_2	ρ
A	0.2	0.2	-0.2	0.3
B	0.2	0.2	-0.2	0.5
C	0.2	0.2	-0.2	0.7
D	0.3	0.2	-0.2	0.3
E	0.3	0.2	-0.2	0.5
F	0.3	0.2	-0.2	0.7

Table 2.2: Simulation results for estimating constant hazard ratios

Setting	t	BIAS	ESD	ASE	ECP
A	1	-0.001	0.129	0.129	0.949
	2	-0.004	0.114	0.114	0.955
	3	-0.004	0.113	0.112	0.954
B	1	-0.002	0.117	0.118	0.960
	2	-0.001	0.107	0.106	0.952
	3	-0.002	0.108	0.107	0.955
C	1	-0.002	0.102	0.105	0.961
	2	0.001	0.097	0.097	0.955
	3	-0.001	0.102	0.102	0.959
D	1	-0.004	0.138	0.138	0.947
	2	-0.005	0.128	0.124	0.946
	3	-0.007	0.129	0.125	0.946
E	1	-0.004	0.127	0.128	0.945
	2	-0.002	0.122	0.118	0.942
	3	-0.003	0.124	0.122	0.955
F	1	-0.007	0.113	0.116	0.948
	2	0.000	0.113	0.110	0.956
	3	-0.002	0.120	0.119	0.958

33% censored first gap times, and 41% censored gap times in subjects with $\{T_{i1} \leq t_1\}$. Under this heavy censoring scenario, three correlation coefficients are used: $\rho = 0.3, 0.5$, and 0.7 (Settings D-F). Table 2.1 lists all six parameter settings. The corresponding results are summarized in Table 2.2. Across all additional settings, we observe similar trends as with Setting B. It is noteworthy that when the correlation between gap times, ρ, increases, the standard error of the estimator decreases.

2.5 Application to liver transplant data

We applied the proposed methods to liver transplant data obtained from the Scientific Registry of Transplant Recipients (SRTR). Our objective is to contrast the

graft failure hazard functions for first and second liver transplants. Graft survival is defined as survival with a functioning graft. Thus, the earliest of transplant failure or death is considered to be the event time. We included adult patients (age $\geq$ 18) that received their first liver transplant between March 1, 2002 and December 31, 2012. Living-donor transplants, transplants to status 1 (acute liver failure) patients, and patients with a MELD exception score were excluded. The end of observation period was December 31, 2012, since death information after that date was possibly incomplete.

In the final data set, we have 31,914 subjects. There are 31,914 primary liver transplants, and 1,566 second transplants. Among the subjects who received primary transplants, 33% are female; and 74% are white. Among the subjects who had repeat transplant, 30% are female; and 73% are white. There are 3,753 graft failures after the first transplant, and 295 graft failures after the second one. For the first transplant, 93% of the graft failures occurred within 5 years; while for the second transplant, 83% of the graft failures occurred within 2 years. For the consideration of enough data, we set $t_1 = 5$ years; in addition, we would evaluate $\theta(t)$ every 30 days until 2 years. The covariates we included can be categorized as being either from the recipient side or the donor side. Recipient-related covariates include age, gender, race, blood type, BMI, calendar year of the transplant, diagnosed diseases, hospitalization information, dialysis, waiting time for a transplant, model for end-stage liver disease (MELD) score components (serum creatinine, bilirubin and INR) and albumin. The donor-related covariates are: serum creatinine, BMI, donor risk index (DRI), warm ischemia time of the organ, whether death was caused by stroke, hypertension status, and diabetes history.

For the first graft failure time T_{i1}, we fit the following proportional hazards model,

Table 2.3: Analysis of SRTR data: Parameter estimates ($p < 0.05$) for the first-liver-transplant Cox model

Covariate		$\widehat{\beta}$	SE($\widehat{\beta}$)	$\widehat{HR}$	p
Race	White	0	0	1	-
	African American	0.332	0.051	1.393	< 0.0001
	Other	0.299	0.149	1.348	0.045
Hepatitis C	Yes	0.266	0.051	1.305	< 0.0001
Non-cholestatic Cirrhosis	Yes	-0.277	0.052	0.758	< 0.0001
Cholestatic Cirrhosis	Yes	-0.181	0.073	0.835	0.013
Hospitalization	No	0	0	1	-
	Yes, ICU	0.160	0.061	1.174	0.009
Age at Transplant		-0.018	0.002	0.982	< 0.0001
Donor BMI	[20,25)	0	0	1	-
	[35,60]	-0.195	0.064	0.823	0.002
log(DRI)		1.196	0.073	3.305	< 0.0001
Donor Diabetes	Yes	0.265	0.052	1.303	< 0.0001
Calendar Year		-0.078	0.009	0.925	< 0.0001

as discussed in the first stage of Section 2.2:

$$\lambda_{i1}(t) = \lambda_0(t) \exp\{\boldsymbol{\beta}' \mathbf{Z}_i\}.$$

The significant parameter estimates are listed in Table 2.3. The censoring model is then fitted using the same covariates. By equation (2.3) in Section 2.2, we estimate θ every 30 days until 2 years. To stabilize the possibly extremely large weights, we cap the weight at 10 (which is approximately the 96th percentile of all maximum within-subject weights). The estimated hazard ratio between transplants ($e^{\widehat{\theta}}$) and its corresponding 95% confidence interval is shown in Figure 2.1. The black solid-dotted line is $\exp\{\widehat{\theta}(t)\}$ estimated at every 30 days until 2 years; the weight is capped at 10; the dashed lines are the estimated 95% confidence interval. As indicated from the figure, the estimated hazard ratio decreases slightly over time before 1 year. After that, the hazard ratio is rather consistent. There is a small bump at around year 1. The reason for this bump is discussed in Section 2.6. We report the final estimates at 2 years in Table 2.4. The estimated hazard ratio is 1.786, with an estimated standard error of 0.107. Thus, the graft failure hazard for repeat liver

Table 2.4: Estimated hazard ratio of first and second liver transplants evaluated at 2 years: Weighted

	Estimate	SE	95% CI	p
$\widehat{\theta}$	0.580	0.060	$(0.463, 0.697)$	< 0.0001
$\exp\{\widehat{\theta}\}$	1.786	0.107	$(1.577, 1.996)$	< 0.0001

Table 2.5: Estimated hazard ratio of first and second liver transplants evaluated at 2 years: Unweighted

	Estimate	SE	95% CI	p
$\widehat{\theta}$	0.409	0.067	$(0.279, 0.540)$	< 0.0001
$\exp\{\widehat{\theta}\}$	1.506	0.100	$(1.309, 1.703)$	< 0.0001

transplantation is estimated to be approximately 80% higher than for primary liver transplantation ($p < 0.0001$), given that the assumption of constant hazard ratio across time holds.

As a comparison, the unweighted estimates are listed in Table 2.5. As shown in the table, without appropriate weighting, the estimated hazard ratio will be artificially underestimated (1.506 vs. 1.786). We could have underestimated the hazard of graft failure after the second liver transplant compared to the first one, if dependent censoring was not accounted for.

Another naive method to estimate the hazard ratio would be to simply stack all data together, from both transplants, and then fit a Cox model using an indicator of re-transplant. This method would be invalid because the issues of identifiability and dependent censoring are not taken into account. The results from this naive analysis are shown in Table 2.6. The estimated hazard ratio is 1.586, which is again biased downward.

Table 2.6: Estimated hazard ratio of first and second liver transplants evaluated at 2 years: the Naive Method

Gap time	$\widehat{\theta}$	SE$(\widehat{\theta})$	$\exp\{\widehat{\theta}\}$	p
1(ref.)	0	0	1	-
2	0.461	0.067	1.586	< 0.0001

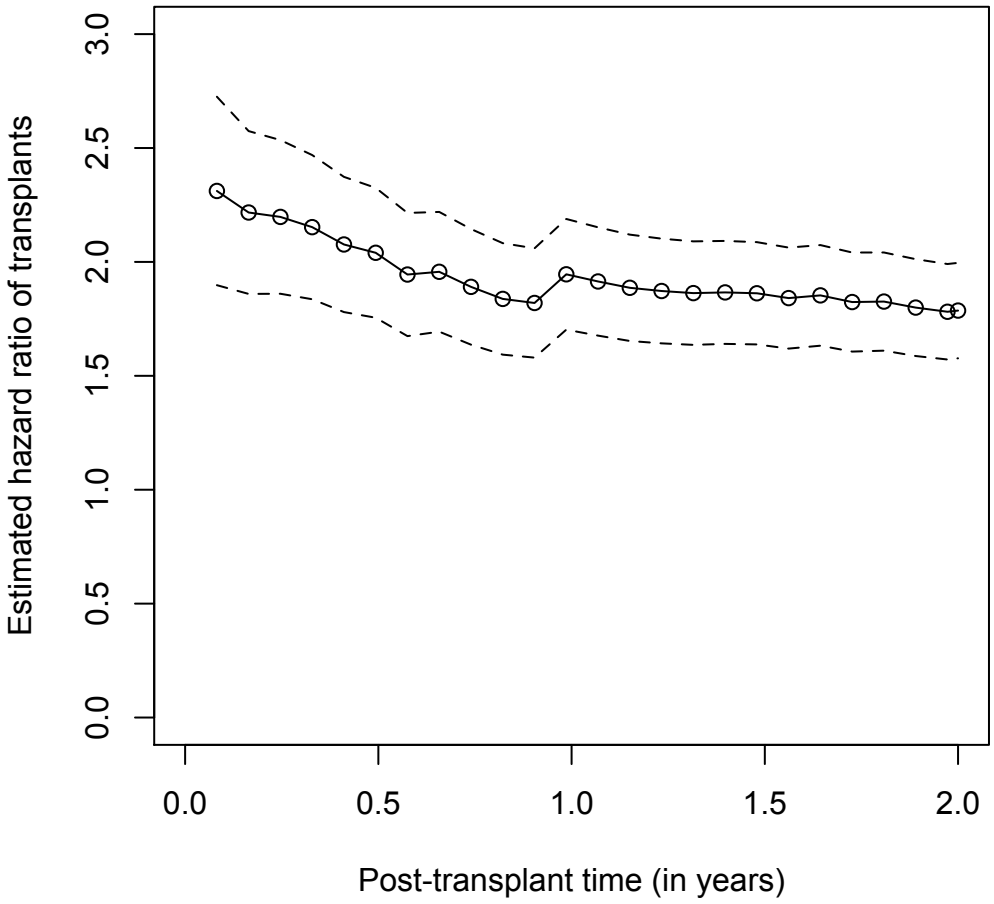

Figure 2.1: Analysis of SRTR data: $\widehat{\theta}(t)$ = estimated hazard ratio of first and second liver transplants

2.6 Discussion

In Chapter I, we have developed semiparametric methods to contrast gap times with respect to survival functions and restricted mean lifetimes. In this chapter, we take a new perspective to contrast the gap times. Specifically, we aim to estimate and compare the hazard ratio between gap times. The concept of hazard ratio is widely used and well accepted by clinicians and doctors; actually, it has almost become the default in many related analyses.

The fundamental assumption of our methods is constant hazard ratio between gap times. The methods consist of two steps. In the first step, we model the hazard function of the first gap time by a proportional hazards model. While in the second step, a weighted estimating equation is solved. The proposed estimator for θ has a nice and clean form: it is closed-form, and is intuitive because it is essentially the observed divided by expected. The estimator can be computed as t-dependent, meaning that there could be various estimates if the hazard ratio varies with t. Normally, the estimates become more centralized and stabilized as t increases, because the accumulated data becomes more abundant. Naturally, any potential gains in stability through using a larger value of t are offset with the turbulence which typically towards the tail of the follow-up period.

We applied the proposed methods to liver transplant data obtained from SRTR. Evaluating at 2 years, we have concluded that the hazard of graft failure after the second transplant is about 80% higher than that after the first transplant. Evaluating at every 30 days until 2 years, we observed a small bump at around 1 year. Although the bump is somewhat inconsequential, we still examined it more carefully. It turned out that one patient has a large weight (≈ 100; capped at 10) at that time. This

patient's first graft survived for 22 days; while the second one was censored alive at 341 days. Due to this patient's particular covariate pattern (The calendar year of transplant for this patient is 2012), the probability that the censoring happens after (22+341) days is quite small. Thus, the resulting weight is large. From this example, we could see that the estimated weight can be very large sometimes. Note that, in the particular setting we consider, capping the weights seems to be the most feasible solution, especially because there does not appear to be an obvious choice for a weight stabilizer.

Our conclusion in the application is in general consistent with what was found in the literature, i.e., the outcomes of the repeat liver transplant are inferior to those of the primary transplant. However, our study is the largest study so far, uses the most up-to-date data, and is statistically valid. In addition, we contrasted the transplants by estimating a hazard ratio, which is not available in most related studies which often rely on results that (in addition to not being properly weighted) were not covariate-adjusted.

The assumption of constant hazard ratio in our report requires that $\lambda_{i1}(t)$ and $\lambda_{i2}(t; t_1)$ have the same span of observation time, so that the hazard ratio could be evaluated at all times. This further determines that t_1 should be less than or equal to $\tau/2$. Thus, the information for $T_1 > t_1$ is not utilized. A possible generalization is to introduce more assumptions, so that all information available could be used. We will focus on developing extensions of the proposed methods in the next chapter.

CHAPTER III

Semiparametric Methods for Contrasting Gap Time Hazard Functions

3.1 Introduction

In many clinical and epidemiological studies, one could experience multiple events during follow-up. There are generally two perspectives to analyze such multiple event data. The 'total time' scale refers to the perspective where one always starts measuring event time from the time of origin, while the 'gap' time measures times between successive events. The research question of interest largely decides which scale is more useful that the other in a given application. In this chapter, we are interested in contrasting times between adjacent events. Thus, we naturally adopt the gap time scale.

Our motivating example arises in the context of liver transplantation. Patients who are diagnosed with end-stage liver disease are often treated with liver transplantation. If graft failure occurs after the first transplant, it is not uncommon for a patient to receive a repeat liver transplant. Currently, the rank of patients on the waiting list is mostly decided by the Model for End-Stage Liver Disease (MELD) score, which is a strong predictor of mortality for patients on the waiting list. However, the ranking system does not take into account whether or not a patient already had a previous liver transplant. This motivates us to contrast the hazard functions

of post-transplant graft survival of first and second transplants (for the ease of presentation; the contrast can be extend to more transplants), adjusting for all other factors. If a second transplant is found to intrinsically result in worse graft survival compared to the first one, hepatologists and liver transplant surgeons may then tend to favor including prior-number-of-transplants when ranking patients on the waiting list. This would be in the interests of optimizing the overall survival for patients given the limited number of available donor livers.

Within-subject correlation between gap times, which is usually present, has led to two most important challenges in analyzing gap time. The first one is known as non-identifiability (Lin et al., 1999; Wang, 1999; Huang, 2002; Schaubel and Cai, 2004a). Specifically, because the support of the first gap time is generally unknown with the presence of censoring, the marginal distributions of the second and subsequent gap times are not identifiable nonparametrically. The second is induced dependent censoring (Visser, 1996; Lin et al., 1999; Huang, 2000). Even if the first gap time is independently censored, the second and subsequent gap times still subject to dependent censoring because of the within-subject correlation.

There are many existing methods in the field of gap time analysis. Some target at estimating the survival functions of gap times nonparametrically. Examples include methods developed by Visser (1996), Wang and Wells (1998), Lin, Sun and Ying (1999), Wang and Chang (1999), Peña, Strawderman and Hollander (2001), van der Lann, Hubbard and Robins (2002), Schaubel and Cai (2004a), and Andrei and Murray (2006). Most other methods seek to model the covariate effects on gap time hazard functions semiparametrically. Prentice, Williams, and Peterson (1981) assumed that the within-subject gap times are independent given the observed covariates, and proposed regression models to relate hazard functions to covariates. Huang

(2002) looked at bivariate response by extending the accelerated failure time model. Huang and Chen (2003) investigated a marginal proportional hazards models for gap times, recognizing that the within-subject correlation may come from random effects. Schaubel and Cai (2004b) developed estimating equations to fit stratified proportional hazards models. Chen, Wang, and Huang (2004) proposed stratified proportional reverse-time hazards models, where the change in hazard across different gap times is estimated by a longitudinal pattern parameter. Strawderman (2005) proposed a semiparametric semi-Markov intensity model, which was later extended for handling correlated gap times by introducing a gamma frailty (Strawderman, 2006). Huang and Liu (2007) proposed a joint frailty model framework to estimate disease recurrences and survival. Clement and Strawderman (2009) modified generalized estimating equations for longitudinal data to estimate the conditional means and variances of the gap times. Du, Jiang, and Wang (2011) proposed a smoothing spline ANOVA frailty model to estimate the gap time hazard nonparametrically.

Despite the richness and diversity of the afore-described literature, almost none of the above-listed methods could be readily adopted or easily generalized to contrast gap times. The few that could have stringent assumptions that are sometimes not applicable to real data. In the first two chapters of this dissertation, we have proposed methods to contrast gap times semiparametrically, with respect to survival functions (Chapter I) and hazard functions (Chapter II). In particular in Chapter II, we have developed methods to contrast the gap-time-specific hazard functions, assuming that the hazard ratio is constant over time. The proposed estimator has a closed form, and it is intuitive to interpret. However, despite its utility, a limitation is that only partial data is utilized in the estimation procedure, due to the conditional nature of the second hazard function combined with the previously described identifiability

issues.

In this chapter, we propose methods to contrast gap times through a time-dependent hazard ratio. Explicitly modeling the dependence of the second gap time hazard on the first gap time enables utilizing the complete data that are available. The assumption specifies the connection between the gap times directly in the model. A function of time is also incorporated, so that a time-dependent hazard ratio can be estimated. On top of the time-dependent hazard ratio, it is often of interest to obtain an overall hazard ratio estimate. We propose a form of weighted average hazard ratio towards the end of Section 3.2.

The remainder of this report is structured as follows. Section 3.2 provides the details of the proposed methods. In Section 3.3, we carried out simulation studies to examine the performance of the estimators under three scenarios. The proposed methods are applied to liver transplant data in Section 3.4. Finally, a discussion is given in Section 3.5.

3.2 Proposed Methods

First, we introduce the necessary notation. Subject is denoted by $i = 1, \cdots, n$. For ease of presentation, we only consider the case of comparing two gap times. Suppose that T_{ij} $(j = 1, 2)$ are the total times of the events, such that T_{i1} is also the first gap time and $\widetilde{T}_{i2} = T_{i2} - T_{i1}$ is the second gap time. The censoring variable is C_i. So the first gap time is potentially censored by C_i, and the second gap time is potentially censored by $\widetilde{C}_{i2} = C_i - T_{i1}$. The covariate vector for subject i is given by $\mathbf{Z}_i$. We define $\tau = \sup\{t : P(C_i \geq t) > 0\}$.

The hazard functions of T_1 and T_2 are given by,

$$\lambda_{i1}(t) = \lim_{\delta \to 0} \delta^{-1} P(t < T_{i1} \leq t + \delta | T_{i1} \geq t, \mathbf{Z}_i)$$

$$\lambda_{i2}(t) = \lim_{\delta \to 0} \delta^{-1} P(t < \widetilde{T}_{i2} \leq t + \delta | \widetilde{T}_{i2} \geq t, \mathbf{Z}_i, T_{i1}).$$

We assume that the hazard functions follow proportional hazards models:

$$\lambda_{i1}(t) = \lambda_0(t) \exp(\boldsymbol{\beta}' \mathbf{Z}_i) \tag{3.1}$$

$$\lambda_{i2}(t) = \lambda_0(t) \exp[\boldsymbol{\beta}' \mathbf{Z}_i + \phi_1 + \phi_2 h_1(T_{i1}) + \phi_3 h_2(t)]. \tag{3.2}$$

This indicates that the correlation between two gap times could be explicitly modeled by including a parametric function of T_{i1} in the model, $h_1(T_{i1})$. In addition, the time-dependent part is modeled by incorporating a function of time, $h_2(t)$. Thus, the subject-specific time-dependent hazard ratio between the two hazard functions is given by,

$$\theta_i(t) = \exp[\phi_1 + \phi_2 h_1(T_{i1}) + \phi_3 h_2(t)].$$

To estimate the population-level time-dependent hazard ratio, we take the expectation of $\theta_i(t)$ with respect to the joint distribution of $\{T_{i1}, \mathbf{Z}_i\}$, i.e.,

$$\widehat{\theta}(t) = \widehat{E}_{\mathbf{Z}} \left\{ \int_0^\tau \exp[\phi_1 + \phi_2 h_1(T_{i1}) + \phi_3 h_2(t)] dG_1(T_{i1} | \mathbf{Z}_i; \tau) \right\}, \tag{3.3}$$

where

$$G_1(t | \mathbf{Z}_i; \tau) = \frac{F_1(t | \mathbf{Z}_i)}{F_1(\tau | \mathbf{Z}_i)},$$

and $F_1(t|\mathbf{Z}_i)$ is the cumulative density function (cdf) of the first gap time. Thus, G_1 is a truncated cdf of T_{i1} restricted on $[0, \tau]$.

It is of interest to not only obtain the average time-dependent hazard ratio, but also to derive an 'overall' average hazard ratio. We thus propose to estimate an overall hazard ratio by further taking an expectation of $\widehat{\theta}(t)$ with respect to the

conditional distribution of $\{\widetilde{T}_{i2}|T_{i1}, \mathbf{Z}_i\}$. The idea is that only the t part needs to be integrated out in $\widehat{\theta}(t)$. Since it is originally built in the second hazard function, it is natural to think it as $\widetilde{T}_{i2}$. Therefore, the final estimate, $\widehat{\theta}$, is given by,

(3.4)

$$\widehat{\theta} = \widehat{E}_{\mathbf{Z}} \left\{ \int_0^\tau \int_0^\tau \exp[\phi_1 + \phi_2 h_1(T_{i1}) + \phi_3 h_2(\widetilde{T}_{i2})] dG_2(\widetilde{T}_{i2}|T_{i1}, \mathbf{Z}_i; \tau) dG_1(T_{i1}|\mathbf{Z}_i; \tau) \right\},$$

where

$$G_2(t|T_{i1}, \mathbf{Z}_i; \tau) = \frac{F_2(t|T_{i1}, \mathbf{Z}_i)}{F_2(\tau|T_{i1}, \mathbf{Z}_i)},$$

and $F_2(t|T_{i1}, \mathbf{Z}_i)$ is the cdf of the second gap time.

It is rarely the case that the parameters and functions are known in (3.3) and (3.4). The parameters ϕ_1, ϕ_2, and ϕ_3 and cdfs F_1 and F_2 could be estimated from the fitted proportional hazards models (3.1) and (3.2). The appropriate functions h_1 and h_2 to be used can be determined by breaking the continuous T_{i1} and t into several interval indicators, fitting a preliminary Cox model with all the indicators, and observe the trend of estimated coefficients over different intervals. The final time-dependent and average hazard ratios are estimated by,

$$\widehat{\theta}(t) = \widehat{E}_{\mathbf{Z}} \left\{ \int_0^\tau \exp[\widehat{\phi}_1 + \widehat{\phi}_2 \widehat{h}_1(T_{i1}) + \widehat{\phi}_3 \widehat{h}_2(t)] d\widehat{G}_1(T_{i1}|\mathbf{Z}_i; \tau) \right\}$$

$$\widehat{\theta} = \widehat{E}_{\mathbf{Z}} \left\{ \int_0^\tau \int_0^\tau \exp[\widehat{\phi}_1 + \widehat{\phi}_2 \widehat{h}_1(T_{i1}) + \widehat{\phi}_3 \widehat{h}_2(\widetilde{T}_{i2})] d\widehat{G}_2(\widetilde{T}_{i2}|T_{i1}, \mathbf{Z}_i; \tau) d\widehat{G}_1(T_{i1}|\mathbf{Z}_i; \tau) \right\},$$

respectively.

In terms of computing the standard error of the proposed estimators, we adopt parametric bootstrapping techniques, exploiting well-established asymptotic properties of the Cox model. The procedure we take here is somewhat different from the usual way of bootstrapping, in order to reduce computational expense. Specifically, for the b-th bootstrap, we re-generate a new set of parameters $\{\widehat{\beta}^{(b)}, \widehat{\phi}_1^{(b)}, \widehat{\phi}_2^{(b)}, \widehat{\phi}_3^{(b)}\}$ from a Multivariate Normal distribution implied by the fitted model. On the other

hand, we generate a new set of baseline hazard functions evaluated at all the original observed unique event times $\{t_1, t_2, \cdots, t_m\}$, written as $\{\widehat{\lambda}_0^{(b)}(t_1), \widehat{\lambda}_0^{(b)}(t_2), \cdots, \widehat{\lambda}_0^{(b)}(t_m)\}$. Then, the statistics $\widehat{\theta}^{(b)}(t)$ and $\widehat{\theta}^{(b)}$ can be re-computed with all the new parameters, assuming that the jump points remain unchanged for the hazard functions. After generating a sufficiently large number of bootstraps (say, B), the standard error of the proposed estimators can be estimated by calculating the sample standard deviation based on $\{\widehat{\theta}^{(1)}(t), \widehat{\theta}^{(2)}(t), \cdots, \widehat{\theta}^{(B)}(t)\}$ and $\{\widehat{\theta}^{(1)}, \widehat{\theta}^{(2)}, \cdots, \widehat{\theta}^{(B)}\}$.

3.3 Simulations

We carried out simulation studies to test how the proposed methods work under different scenarios. We simulate $n = 500$ subjects. A baseline covariate, Z_i, is generated for each subject from a Bernoulli(0.5) distribution. The first and second gap times are generated from the following two proportional hazards models,

$$\lambda_{i1}(t) = \lambda_0(t) \exp(\beta Z_i) \tag{3.5}$$

$$\lambda_{i2}(t) = \lambda_0(t) \exp(\beta Z_i + \phi_1 + \phi_2 T_{i1} + \phi_3 t), \tag{3.6}$$

respectively, where $\lambda_0(t) = 0.4$, $\beta = \log(1.5)$, $\phi_1 = 0.2$, $\phi_2 = -0.2$, and $\phi_3 = 0.2$. The censoring variable, C_i, follows a Uniform distribution on $(0, 10)$. Under this data configuration, approximately 78% of the first gap times are observed; and 62% of the second gap times are observed.

Based on the proposed methods, we estimated the population-level time-dependent hazard ratio by

$$\widehat{\theta}(t) = \widehat{E}_Z \left\{ \int_0^\tau \exp(\widehat{\phi}_1 + \widehat{\phi}_2 T_{i1} + \widehat{\phi}_3 t) d\widehat{G}_1(T_{i1} | \mathbf{Z}_i; \tau) \right\}, \tag{3.7}$$

where $\widehat{\phi}_1, \widehat{\phi}_2, \widehat{\phi}_3, \widehat{G}_1$ were estimated based on Cox models (3.5) and (3.6), and E_Z was estimated empirically by taking an average across all subjects.

Table 3.1: Simulation results for estimating time-dependent and average hazard ratios: $\phi_3 = 0.2$

t	$\theta(t)$	$\widehat{\theta}(t)$	BIAS	BSE	ESD	95% ECP
0	0.985	0.983	-0.002	0.118	0.112	0.957
0.5	1.089	1.089	-0.001	0.106	0.100	0.959
1	1.203	1.205	0.002	0.100	0.095	0.957
1.5	1.330	1.337	0.007	0.108	0.104	0.959
2	1.470	1.484	0.014	0.136	0.134	0.955
2.5	1.624	1.649	0.025	0.185	0.183	0.959
3	1.795	1.834	0.039	0.253	0.250	0.960
Average	**1.260**	**1.268**	**0.008**	**0.090**	**0.092**	**0.941**

The average hazard ratio was then estimated by,

$$\widehat{\theta} = \widehat{E}_Z \left\{ \int_0^\tau \int_0^\tau \exp(\widehat{\phi}_1 + \widehat{\phi}_2 T_{i1} + \widehat{\phi}_3 \widetilde{T}_{i2}) d\widehat{G}_2(\widetilde{T}_{i2}|T_{i1}, Z_i; \tau) d\widehat{G}_1(T_{i1}|Z_i; \tau) \right\}.$$

We set $\tau = 3$ and estimate $\theta(t)$ for $t = 0, 0.5, 1.0, \ldots, 2.5, 3.0$. We ran 1,000 replicates per configuration, each with $B = 50$ bootstrap samples. The estimated time-dependent hazard ratios are shown in Table 3.1. The last row shows the overall average hazard ratio. As seen from the table, biases are small, especially at earlier time points. The increasing trend in bias as time goes on is due to the nature of the data configuration. Specifically, we have $\phi_3 > 0$ in (3.6). After some simple deductions, it can be shown that the bias of $\widehat{\theta}(t)$ at a later time point would be the bias at an earlier time point multiplied by a factor that converges to $\exp(\phi_3)$ (which will be > 1). The bootstrapped standard errors (BSE) and the empirical standard deviations (ESD) are close across all time points, indicating that the bootstrapped standard errors are quite accurate. The 95% estimated cover probability (ECP) is around 0.95. The coverage probability is slightly bigger than 0.95 (0.972) for the average hazard ratio, suggesting slightly conservative inference.

In addition, we have simulated another set of data configuration, where the time-dependent hazard ratio decreases over time ($\phi_3 < 0$). We generate T_1 and T_2 from

Table 3.2: Simulation results for estimating time-dependent and average hazard ratios: $\phi_3 = -0.75$

t	$\theta(t)$	$\widehat{\theta}(t)$	BIAS	BSE	ESD	95% ECP
0	0.921	0.932	0.011	0.147	0.138	0.959
1	0.547	0.549	0.002	0.049	0.044	0.960
2	0.404	0.405	0.001	0.043	0.041	0.950
3	0.325	0.328	0.003	0.045	0.044	0.945
4	0.275	0.278	0.003	0.047	0.046	0.945
5	0.240	0.244	0.004	0.049	0.047	0.947
Average	**0.541**	**0.548**	**0.007**	**0.067**	**0.053**	**0.975**

the following models:

$$\lambda_{i1}(t) = \lambda_0(t)\exp(\beta Z_i)$$

$$\lambda_{i2}(t) = \lambda_0(t)\exp[\beta Z_i + \phi_1 + \phi_2 T_{i1} + \phi_3 \log(t+1)],$$

where all the parameter values are the same as before except that $\phi_3 = -0.75$.

We set $\tau = 5$ in this data configuration, a more appropriate choice given the particular distribution of the gap times. The time-dependent hazard ratio is estimated from $t = 0$ until $t = 5$ with increments of 1. The results are shown in Table 3.2. Again, the biases are very small, while BSE and ESD match quite well. We would imagine to see a decrease trend in bias in Table 3.2 for an opposite reason for the increase in Table 3.1. Indeed, we do observe a decrease from $t = 0$ to 1, but then the bias slightly increased for the later time points. This is because as t approximates τ, we have thinner data towards the tail, and thus slightly more bias.

The last scenario we considered is associated with a even stronger hazard decrease trend over time. In particular, we set $\phi_3 = -1$. The models for generating the gap times remain the same as in the second scenario. The results are shown in Table 3.3. Again, we observe similar patterns in results as in the previous simulations. To summarize, the proposed estimators and standard errors work well under the scenarios we have considered.

Table 3.3: Simulation results for estimating time-dependent and average hazard ratios: $\phi_3 = -1$

t	$\theta(t)$	$\widehat{\theta}(t)$	BIAS	BSE	ESD	95% ECP
0	0.921	0.932	0.011	0.151	0.144	0.952
1	0.460	0.463	0.003	0.043	0.042	0.948
2	0.307	0.310	0.003	0.034	0.033	0.951
3	0.230	0.233	0.003	0.034	0.033	0.955
4	0.184	0.188	0.004	0.034	0.032	0.955
5	0.153	0.157	0.004	0.033	0.031	0.956
Average	**0.482**	**0.489**	**0.005**	**0.067**	**0.057**	**0.972**

3.4 Application to liver transplant data

We applied the proposed methods to liver transplant data obtained from the Scientific Registry of Transplant Recipients (SRTR). The event of interest is the earliest of graft failure or death. We aim to estimate the hazard functions of post-transplant event for the first and second transplant and compute the hazard ratio between the two, both the time-dependent and average ones. Patients who received their first liver transplant between March 1, 2002 and December 31, 2011 and aged ≥ 18 at the time of their first transplant are included in the analysis. In addition, we exclude patients who had at least one of the following: living-donor transplants, transplants to status 1 (acute liver failure), and MELD exceptions. We set the end of observation date to be December 31, 2012, since death information beyond that date is possibly incomplete.

After applying the inclusion and exclusion criteria, the resulting cohort is composed of 31,914 subjects. Among those, there are 1,566 patients who received a second liver transplant. The number of events after the first transplant is 3,753; the number after the second one is 295. It is the same data set as we analyzed in Chapter II. The covariates ($\mathbf{Z}_i$) are either recipient-related or donor-related. Recipient-related covariates include age, gender, race, blood type, BMI, calendar year of the

transplant, diagnosed diseases, hospitalization information, dialysis, waiting time for a transplant, model for end-stage liver disease (MELD) score components (serum creatinine, bilirubin and INR) and albumin. Donor-related covariates are: serum creatinine, BMI, donor risk index (DRI), warm ischemia time of the organ, whether death was caused by stoke, hypertension status, and diabetes history. The following Cox model is fitted:

$$\lambda_{ij}(t) = \lambda_0(t)\exp\{\boldsymbol{\beta}'\mathbf{Z}_i + \phi_1 I(j=2) + \phi_2 I(j=2)h_1(T_{i1}) + \phi_3 I(j=2)h_2(t)\}, j = 1, 2.$$

In order to determine the functional forms of h_1 and h_2, we break the continuous covariates T_{i1} and t into categorical versions of variables, and fit the above model. After plotting and observing the trend of estimated coefficients across different intervals, we choose the following spline functions:

$$h_1(T_{i1}) = h_{11}T_{i1} + h_{12}(T_{i1} - 365)_+$$

$$h_2(t) = h_{21}t + h_{22}(t - 180)_+ + h_{23}(t - 365)_+ + h_{24}(t - 1095)_+,$$

where $(x - t)_+ \equiv (x - t)I(x > t)$.

We set $\tau = 5$ years. Applying the proposed methods, we estimate the hazard ratio between the two transplants at every months from 0 to 5 years; an average hazard ratio is also obtained. Due to the fact that we have an enormous amount of observations in the final data set, and the time-dependent nature of the modeling essentially expands the data set even more considerably, we are faced with high computational time. Therefore, we select a 'case-control' sample from the data set and estimate the statistics and the corresponding bootstrapped standard errors. The sample is drawn through the following steps. First, records of patients who had a repeat transplant are sampled with probability 1, so that the information of the correlation between gap times remains as complete as possible. Then, all 'event'

records for the first transplant are sampled with probability 1, i.e., we sample all the 'cases'. Finally, we sample from the remaining records randomly with probability 0.2, with a sampling weight of 5 applied. After one-time draw, the resulting sample has 11,101 observations.

The parameter estimates from the Cox model, using the whole data set, are listed in Table 3.4. Note that only the significant ones are shown here. The estimated hazard ratios are summarized in Table 3.5. Not every month is shown because the table would otherwise be too lengthy. $\widehat{\theta}_s(t)$ is the hazard ratio at time t estimated from sample, and $\widehat{\theta}_w(t)$ is the one estimated from the whole data set. As seen, $\widehat{\theta}_s(t)$ and $\widehat{\theta}_w(t)$ are very similar at all time points, indicating that the sample represents the whole data set well. The time-dependent hazard ratio between first and second transplants is estimated to be 2.176 ($p < 0.0001$) at the beginning, i.e., the hazard of graft failure right after the second liver transplant is 2.176 times larger than that right after the first transplant. The hazard ratio starts to decrease rapidly during the first 6 months after transplant, and it reaches below 1 at 6th month. The hazard ratio seems to gradually stabilize at around slightly below 1 after 6 months. From a statistical significance point of view, the hazard ratio becomes not significantly different from the null at 4 months. Overall, the hazard of graft failure after the second transplant is estimated to be 1.388 ($p < 0.0001$) times bigger than that after the first one, while the time-dependent ones suggest that the larger hazard of the second transplant mainly results from the first four months after transplant.

Figure 3.1 shows the estimated time-dependent hazard ratios at every month from 0 to 5 years. The black solid-dotted line is $\exp\{\widehat{\theta}(t)\}$ estimated at every month until 5 years; the dashed lines are the 95% confidence interval estimated through bootstrapping. The confidence interval seems to fluctuate a bit from 6 months to 5

Table 3.4: Analysis of SRTR data: Parameter estimates ($p < 0.05$) for Cox model

Covariate		$\widehat{\beta}$	SE($\widehat{\beta}$)	$\widehat{HR}$	p
Race	White	0	0	1	-
	African American	0.313	0.049	1.368	< 0.0001
	Other	0.299	0.143	1.349	0.037
Hepatitis C	Yes	0.221	0.049	1.247	< 0.0001
Non-cholestatic Cirrhosis	Yes	-0.300	0.049	0.741	< 0.0001
Cholestatic Cirrhosis	Yes	-0.218	0.070	0.804	0.002
Metastatic Disease	Yes	-0.242	0.110	0.785	0.028
Hospitalization	No	0	0	1	-
	Yes, ICU	0.186	0.058	1.204	0.002
Age at Transplant		-0.018	0.002	0.982	< 0.0001
Donor BMI	[20,25)	0	0	1	-
	[35,60]	-0.179	0.061	0.836	0.003
log(DRI)		1.112	0.070	3.042	< 0.0001
Donor Diabetes	Yes	0.257	0.050	1.292	< 0.0001
Calendar Year		-0.074	0.009	0.928	< 0.0001
Re-transplant	Yes	0.826	0.121	2.284	< 0.0001
t^*		-1.817	0.505	0.162	0.0003
$(t > 180/365)(t - 180/365)^*$		2.126	1.015	8.383	0.036

* t is in years.

years, largely due to the fact that the hazard ratio is computed instantaneously, and the standard errors largely depend on the number of events and number at risk at that particular time point.

3.5 Discussion

In this report, we propose methods that are generated from the idea in Chapter II to estimate and contrast time-dependent hazard functions of gap times. The proportional hazards modeling framework we develop is of a flexible form, with both the connection between gap times and how hazard function changes over time explicitly specified. We propose to estimate the time-dependent hazard ratio and the overall average hazard ratio by taking expectations to appropriate joint distributions of gap times. Therefore, the estimators naturally place a larger weight for the observations where events are abundant. Simulations studies have demonstrated that the proposed estimators are efficient in recovering the true hazard ratios.

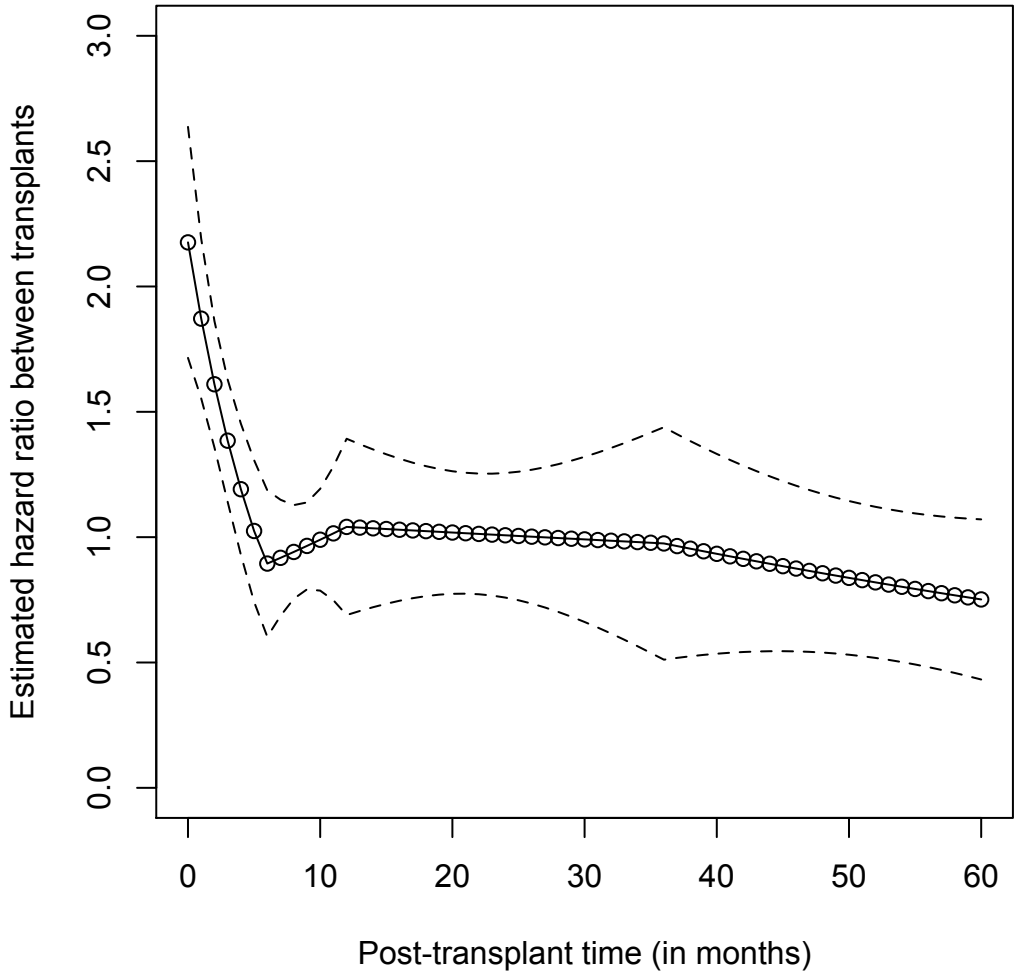

Figure 3.1: Analysis of SRTR data: $\widehat{\theta}(t)$ = estimated time-dependent hazard ratio of first and second liver transplants

Table 3.5: Analysis of SRTR data: Time-dependent and average hazard ratios between liver transplants

t (months)	$\widehat{\theta}_w(t)$	$\widehat{\theta}_s(t)$	BSE	p
0	2.070	2.176	0.235	< 0.0001
1	1.779	1.872	0.162	< 0.0001
2	1.529	1.610	0.129	< 0.0001
3	1.314	1.385	0.124	0.0003
4	1.129	1.191	0.133	0.117
5	0.970	1.025	0.144	0.870
6	0.847	0.895	0.149	0.516
12	0.988	1.041	0.179	0.827
18	0.970	1.024	0.132	0.864
24	0.952	1.007	0.127	0.960
30	0.934	0.991	0.168	0.961
36	0.917	0.975	0.237	0.924
42	0.857	0.914	0.190	0.678
48	0.802	0.856	0.161	0.416
54	0.750	0.802	0.153	0.250
60	0.701	0.752	0.163	0.190
Average	**1.464**	**1.388**	**0.089**	< 0.0001

Although the methods proposed can be viewed as a generalization of the methods in Chapter II, they differ considerably. The method we propose here could utilize all data that are available, and it enables one to estimate the hazard ratio between gap times as a function of time. While in Chapter II, only partial data could be utilized due to the conditional definition of the second hazard function; in addition, the underlying assumption was that the hazard ratio is constant over time. It is true that we can change the upper bound of integration in the estimator to compute a time-dependent-like hazard ratio. However, such hazard ratio has very different interpretation with the hazard ratio we estimate in this chapter. Specifically, the former one (in Chapter II) is a 'cumulative' one due to the interpretation of the estimator (observed divided by expected), while the latter one is more of an 'instantaneous' nature because it averages the subject-specific hazard ratios at a particular time point. Therefore, the hazard ratios that are estimated from Chapter II and Chapter III cannot be simply compared. In fact, they contain different information

and are useful in different ways.

The proposed methods are applied to the same data set that is analyzed in Chapter II. The hazard of graft failure after the second transplant is significantly bigger than that after the first one until the 4th month. There is no evidence against a null hazard ratio (= 1) from 4 months to 5 years. Overall, the average hazard of graft failure for the second transplant is still significantly bigger. We conclude that the difference in hazard functions between gap times is prominent in a very short period of time right after transplant.

APPENDICES

APPENDIX A

Supplementary Materials for Chapter I

In this appendix, we prove Theorems 1.1, 1.2 and 1.3 from Section 1.3, describe the additional simulations that were carried out, and list the parameter estimates for the gap-time-specific models. To begin we review the notation.

A.1 Notation

i = subject $(i = 1, \cdots, n)$

T_{i1} = first gap time

$\widetilde{T}_{i2}$ = second gap time

C_i = censoring time for T_{i1}

$\widetilde{C}_{i2} = C_i - T_{i1}$ = censoring time for $\widetilde{T}_{i2}$

L = mean survival time cap

$\tau_1 = \sup\{t : P(C_i > t) > 0\}$

$\tau_2 = \sup\{t : P(\widetilde{C}_{i2} > t) > 0\}$

$\mathbf{Z}_i$ = time-constant covariate vector

$\mathbf{W}_i = (\mathbf{Z}_i', \boldsymbol{f}(T_{i1})')'$, where $\boldsymbol{f}$ is a parametric possibly vector valued function

$\lambda_{i1}(t|\mathbf{Z}_i) = \lambda_{01}(t)\exp\{\boldsymbol{\beta}_1'\mathbf{Z}_i\}$

$\lambda_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1) = \lambda_{02}(t; \tau_1)\exp\{\boldsymbol{\beta}_2'\mathbf{Z}_i + \boldsymbol{\phi}_2'\boldsymbol{f}(T_{i1})\}$

$\boldsymbol{\theta}_2 = (\boldsymbol{\beta}_2', \boldsymbol{\phi}_2')'$

$S_{i1}(t|\mathbf{Z}_i) = P(T_{i1} > t|\mathbf{Z}_i)$

$S_{i2}(t|\mathbf{Z_i}, T_{i1}; \tau_1) = P(\widetilde{T}_{i2} > t|\mathbf{Z}_i, T_{i1}, T_{i1} \leq \tau_1)$

$\widehat{S}_{i1}(t|\mathbf{Z}_i) = \exp\{-\widehat{\Lambda}_{01}(t)\exp\{\widehat{\boldsymbol{\beta}}_1'\mathbf{Z}_i\}\}$

$\widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1) = \exp\{-\widehat{\Lambda}_{02}(t;\tau_1)\exp\{\widehat{\boldsymbol{\theta}}_2'\mathbf{W}_i\}\}$

$T_{i1}^m = m$'th imputed value of T_{i1}, for $m = 1, \ldots, M$

$M =$ number of imputations

$\mathbf{W}_i^m = (\mathbf{Z}_i', \boldsymbol{f}(T_{i1}^m)')'$

$\widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) = \exp\{-\widehat{\Lambda}_{02}(t;\tau_1)\exp\{\widehat{\boldsymbol{\theta}}_2'\mathbf{W}_i^m\}\}$

$\widehat{\delta}_i^m(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) = \widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m; \tau_1) - \widehat{S}_{i1}(t|\mathbf{Z}_i)$

$\widehat{\Delta}_i^m(L|\mathbf{Z}_i, T_{i1}^m; \tau_1) = \int_0^L \widehat{\delta}_i^m(t|\mathbf{Z}_i, T_{i1}^m; \tau_1)dt$

$N_{i1}(t) = I\{T_{i1} \leq t \wedge C_i\}$

$N_{\bullet 1}(\tau_1) = \sum_{i=1}^n N_{i1}(\tau_1) = \sum_{i=1}^n I\{T_{i1} \leq \tau_1 \wedge C_i\}$

$N_{i1}^m(t) = I\{T_{i1}^m \leq t\}$

$N_{\bullet 1}^m(\tau_1) = \sum_{i=1}^n I\{T_{i1}^m \leq \tau_1\}$

$\widehat{S}_1^m(t;\tau_1) = N_{\bullet 1}^m(\tau_1)^{-1}\sum_{i=1}^n N_{i1}^m(\tau_1)\widehat{S}_{i1}(t|\mathbf{Z}_i)$

$\widehat{S}_2^m(t;\tau_1) = N_{\bullet 1}^m(\tau_1)^{-1}\sum_{i=1}^n N_{i1}^m(\tau_1)\widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m;\tau_1)$

$\widehat{\delta}^m(t;\tau_1) = \widehat{S}_2^m(t;\tau_1) - \widehat{S}_1^m(t;\tau_1)$

$\widehat{\Delta}^m(L;\tau_1) = \int_0^L \widehat{\delta}^m(t;\tau_1)dt$

$\widehat{S}_1(t;\tau_1) = M^{-1}\sum_{m=1}^M \widehat{S}_1^m(t;\tau_1)$

$\widehat{S}_2(t;\tau_1) = M^{-1}\sum_{m=1}^M \widehat{S}_2^m(t;\tau_1)$

$\widehat{\delta}(t;\tau_1) = \widehat{S}_2(t;\tau_1) - \widehat{S}_1(t;\tau_1)$

$\widehat{\Delta}(L;\tau_1) = \int_0^L \widehat{\delta}(t;\tau_1)dt$

$\widetilde{N}_{i2}(t) = I\{\widetilde{T}_{i2} \leq t \wedge \widetilde{C}_{i2}, T_{i1} \leq \tau_1 \wedge C_i\}$

$Y_{i1}(t) = I\{T_{i1} \wedge C_i \geq t\}$

$\widetilde{Y}_{i2}(t) = I\{\widetilde{T}_{i2} \wedge \widetilde{C}_{i2} \geq t, T_{i1} \leq \tau_1 \wedge C_i\}$

$M_{i1}(t) = N_{i1}(t) - \int_0^t \lambda_{i1}(u) Y_{i1}(u) du$

$\widetilde{M}_{i2}(t) = \widetilde{N}_{i2}(t) - \int_0^t \lambda_{i2}(u;\tau_1) \widetilde{Y}_{i2}(u) du$

Next, we list the assumed conditions underlying our proofs.

A.2 Regularity Conditions

We assume the following regularity conditions for $i = 1, \cdots, n$, $0 \le s \le \tau_1$, $0 \le u \le \tau_2$:

(a) $\{N_{i1}(\cdot), \widetilde{N}_{i2}(\cdot), Y_{i1}(\cdot), \widetilde{Y}_{i2}(\cdot), \mathbf{Z}_i\}$ are independent and identically distributed;

(b) $E[Y_{i1}(s)] > 0$ and $E[\widetilde{Y}_{i2}(u)] > 0$;

(c) elements of $\mathbf{Z}_i$ are bounded almost surely;

(d) $\Lambda_{01}(s) < \infty$ and $\Lambda_{02}(u;\tau_1) < \infty$;

(e) positive-definiteness of the following matrices:

$$\begin{aligned}
\boldsymbol{\Sigma}_1(\boldsymbol{\beta}) &= E\left[\int_0^{\tau_1} \left\{ \frac{\mathbf{s}_1^{(2)}(t,\boldsymbol{\beta}_1)}{s_1^{(0)}(t,\boldsymbol{\beta}_1)} - \overline{\boldsymbol{z}}_1(t,\boldsymbol{\beta}_1)^{\otimes 2} \right\} dN_{i1}(t)\right], \\
\boldsymbol{\Sigma}_2(\boldsymbol{\theta}) &= E\left[\int_0^{\tau_1} \pi_{i1}(u,\mathbf{Z}_i;\tau_1) du \int_0^{\tau_2} \left\{ \frac{\mathbf{s}_2^{(2)}(t,\boldsymbol{\theta}_2)}{s_2^{(0)}(t,\boldsymbol{\theta}_2)} - \overline{\boldsymbol{z}}_2(t,\boldsymbol{\theta}_2)^{\otimes 2} \right\} d\widetilde{N}_{i2}(t)\right].
\end{aligned}$$

A.3 Proofs of Theorems 1.1, 1.2 and 1.3

We first provide expressions for quantities referred to in Theorem 1.1 but not explicitly listed in Section 1.3.

A.3.1 Continuation of Theorem 1.1

The following quantities are pertinent to Theorem 1.1:

$$\begin{aligned}
\phi_{i1}^m(t) &= \mathbf{b}_m(t)' \int_0^{\tau_2} [\mathbf{W}_i^m - \overline{\boldsymbol{z}}_2(t,\boldsymbol{\theta}_2)] d\widetilde{M}_{i2}(t) \\
&\quad - E[S_{i2}(t|\mathbf{W}_i;\tau_1) \exp(\boldsymbol{\theta}_2'\mathbf{W}_i) | T_{i1}^m \le \tau_1] \int_0^t s_2^{(0)}(u,\boldsymbol{\theta}_2)^{-1} d\widetilde{M}_{i2}(u),
\end{aligned}$$

where $S_{i2}(t|\mathbf{W}_i|T_{i1} \leq \tau_1) = S_{i2}(t|\mathbf{Z}_i, T_{i1}; \tau_1)$ by definition (since $\mathbf{W}_i$ is a function of $\mathbf{Z}_i$ and T_{i1}) and with

$$\begin{aligned}\mathbf{b}_m(t) &= \boldsymbol{\Sigma}_2(\boldsymbol{\theta})^{-1} \int_0^t \{\overline{\boldsymbol{z}}_2(u, \boldsymbol{\theta}_2) E[S_{i2}(t|\mathbf{W}_i; \tau_1) \exp(\boldsymbol{\theta}_2' \mathbf{W}_i) | T_{i1}^m \leq \tau_1] \\ &\quad - E[\mathbf{W}_i S_{i2}(t|\mathbf{W}_i; \tau_1) \exp(\boldsymbol{\theta}_2' \mathbf{W}_i) | T_{i1}^m \leq \tau_1]\} d\Lambda_{02}(u; \tau_1),\end{aligned}$$

and $\boldsymbol{\Sigma}_2(\boldsymbol{\theta})$ is as defined in Condition (e).

Next, we provide the remaining expressions alluded to in Theorem 1.2.

A.3.2 Continuation of Theorem 1.2

The following quantities are pertinent to Theorem 1.2: Specifically,

$$\begin{aligned}\varphi_{i3}(t) =& M^{-1} \sum_{m=1}^{M} \left\{ \mathbf{a}(t)' \int_0^{\tau_1} [\mathbf{Z}_i - \overline{\boldsymbol{z}}_1(t, \boldsymbol{\beta}_1)] dM_{i1}(t) - E[S_1(t|\mathbf{Z}_i) \exp(\boldsymbol{\beta}_1' \mathbf{Z}_i)] \right. \\ & \left. \times \int_0^t s_1^{(0)}(u, \boldsymbol{\beta}_1)^{-1} dM_{i1}(u) + P(T_{i1} \leq \tau_1)^{-1} N_{i1}^m(\tau_1) [S_{i1}(t|\mathbf{Z}_i) - S_1(t; \tau_1)] \right\},\end{aligned}$$

$$\varphi_{i4}(L) = \int_0^L \varphi_{i3}(t) dt,$$

where

$$\mathbf{a}(t) = \boldsymbol{\Sigma}_1(\boldsymbol{\beta})^{-1} \int_0^t \{\overline{\boldsymbol{z}}_1(u, \boldsymbol{\beta}_1) E[S_1(t|\mathbf{Z}_i) \exp(\boldsymbol{\beta}_1' \mathbf{Z}_i)] - E[\mathbf{Z}_i S_1(t|\mathbf{Z}_i) \exp(\boldsymbol{\beta}_1' \mathbf{Z}_i)]\} d\Lambda_{01}(u),$$

and $\boldsymbol{\Sigma}_1(\boldsymbol{\beta})$ is as defined in Condition (e).

We next prove Theorem 1.1. Before that, we take a closer look at the imputation mechanism and prove some useful results about the relationship between the imputed and true distributions of T_{i1}.

The following lemma is essential to the proofs of Theorems 1.1 to 1.3.

A.3.3 Lemma and its Proof

LEMMA A.1: *The distribution of T_{i1}^m is equivalent to the distribution of T_{i1}. Hence, $N_{\bullet 1}^m(\tau_1)/n$ converges in probability to $P(T_{i1} \leq \tau_1)$.*

Proof of Lemma: To prove this lemma, we recall the imputation procedure. For subjects with first gap time observed, we let $T_{i1}^m = T_{i1}$; for subjects with first gap time censored, T_{i1}^m are generated from the estimated truncated distribution $\widehat{P}(T_{i1}^m > t|C_i, T_{i1} > C_i, \mathbf{Z}_i)$=$\widehat{P}(T_{i1} > t|C_i, T_{i1} > C_i, \mathbf{Z}_i) = \widehat{S}_{i1}(t|\mathbf{Z}_i)/\widehat{S}_{i1}(C_i|\mathbf{Z}_i)$.

The true underlying distribution of T_{i1}^m when $T_{i1} > C_i$ is given by,

$$P(T_{i1}^m > t|C_i, T_{i1} > C_i, \mathbf{Z}_i) = \frac{S_{i1}(t|\mathbf{Z}_i)}{S_{i1}(C_i|\mathbf{Z}_i)} = \frac{P(T_{i1} > t|\mathbf{Z}_i)}{P(T_{i1} > C_i|C_i, \mathbf{Z}_i)}. \tag{A.1}$$

Then, the proof of the first part of Lemma A.1 follows by splitting $P(T_{i1}^m > t|\mathbf{Z}_i)$ into two parts based on whether $T_{i1} \leq C_i$ or $T_{i1} > C_i$. Assume C_i has a probability density function f_{C_i}, we have

$$\begin{aligned}
&P(T_{i1}^m > t|\mathbf{Z}_i) \\
&\quad = P(T_{i1} > t, T_{i1} \leq C_i|\mathbf{Z}_i) + P(T_{i1}^m > t, T_{i1} > C_i|\mathbf{Z}_i) \\
&\quad = \int_t^\infty P(t < T_{i1} \leq c|\mathbf{Z}_i) f_{C_i}(c)dc + \int_0^t P(T_{i1}^m > t, T_{i1} > C_i|C_i = c, \mathbf{Z}_i) f_{C_i}(c)dc \\
&\qquad + \int_t^\infty P(T_{i1} > c|\mathbf{Z}_i) f_{C_i}(c)dc \\
&\quad = \int_t^\infty P(T_{i1} > t|\mathbf{Z}_i) f_{C_i}(c)dc \\
&\qquad + \int_0^t P(T_{i1}^m > t|C_i = c, T_{i1} > C_i, \mathbf{Z}_i) \times P(T_{i1} > C_i|C_i = c, \mathbf{Z}_i) f_{C_i}(c)dc \\
&\overset{(A.1)}{=} \int_t^\infty P(T_{i1} > t|\mathbf{Z}_i) f_{C_i}(c)dc + \int_0^t \frac{P(T_{i1} > t|\mathbf{Z}_i)}{P(T_{i1} > c|\mathbf{Z}_i)} \times P(T_{i1} > c|\mathbf{Z}_i) f_{C_i}(c)dc \\
&\quad = \int_0^\infty P(T_{i1} > t|\mathbf{Z}_i) f_{C_i}(c)dc \\
&\quad = P(T_{i1} > t|\mathbf{Z}_i).
\end{aligned}$$

Taking expectation with respect to $\mathbf{Z}$ to both sides of the equation, we can argue that,

$$P(T_{i1}^m > t) = P(T_{i1} > t).$$

The second part of the Lemma can be readily derived. Since $N^m_{\bullet 1}(\tau_1) = \sum_{i=1}^n I\{T^m_{i1} \leq \tau_1\}$, we have,

$$\frac{N^m_{\bullet 1}(\tau_1)}{n} = \frac{1}{n}\sum_{i=1}^{n} I\{T^m_{i1} \leq \tau_1\} \xrightarrow{p} P(T^m_{i1} \leq \tau_1) = P(T_{i1} \leq \tau_1),$$

with convergence in probability from the Weak Law of Large Numbers (WLLN).

To prove Theorem 1.1, we first prove consistency, then derive each component of the pertinent linear representations.

A.3.4 Consistency

By the WLLN and Lemma A.1, we have

$$\begin{aligned}
&\widehat{S}_2(t;\tau_1) - S_2(t;\tau_1) \\
&\quad = M^{-1}\sum_{m=1}^{M} N^m_{\bullet 1}(\tau_1)^{-1}\sum_{i=1}^{n} N^m_{i1}(\tau_1)[\widehat{S}_{i2}(t|\mathbf{Z}_i, T^m_{i1};\tau_1) - S_2(t;\tau_1)] \\
&\quad = M^{-1}\sum_{m=1}^{M} \frac{n}{N^m_{\bullet 1}(\tau_1)} n^{-1}\sum_{i=1}^{n} N^m_{i1}(\tau_1)[\widehat{S}_{i2}(t|\mathbf{Z}_i, T^m_{i1};\tau_1) - S_2(t;\tau_1)] \\
&\quad \xrightarrow{p} \frac{1}{M}\sum_{m=1}^{M} P(T_{i1} \leq \tau_1)^{-1} P(T_{i1} \leq \tau_1) E\{[\widehat{S}_{i2}(t|\mathbf{Z}_i, T^m_{i1};\tau_1) - S_2(t;\tau_1)]|T^m_{i1} \leq \tau_1\} \\
&\quad \xrightarrow{p} 0.
\end{aligned}$$

Using continuity, we can also argue

$$\widehat{\mu}_2(L;\tau_1) - \mu_2(L;\tau_1) = \int_0^L [\widehat{S}_2(t;\tau_1) - S_2(t;\tau_1)]dt \xrightarrow{p} 0.$$

Thus, the consistency of $\widehat{S}_2(t;\tau_1)$ and $\widehat{\mu}_2(L;\tau_1)$ for their respective target values is proved.

A.3.5 Linear Representations

We can decompose the target quantity as follows,

$$
\begin{aligned}
& n^{1/2}[\widehat{S}_2(t;\tau_1) - S_2(t;\tau_1)] \\
& \quad = n^{1/2} M^{-1} \sum_{m=1}^{M} N_{\bullet 1}^m(\tau_1)^{-1} \sum_{i=1}^{n} N_{i1}^m(\tau_1)[\widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m;\tau_1) - S_2(t;\tau_1)] \\
& \quad = n^{1/2} M^{-1} \sum_{m=1}^{M} N_{\bullet 1}^m(\tau_1)^{-1} \sum_{i=1}^{n} N_{i1}^m(\tau_1)[\widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m;\tau_1) - S_{i2}(t|\mathbf{Z}_i, T_{i1}^m;\tau_1)] \qquad \text{(A.2)} \\
& \qquad + n^{1/2} M^{-1} \sum_{m=1}^{M} N_{\bullet 1}^m(\tau_1)^{-1} \sum_{i=1}^{n} N_{i1}^m(\tau_1)[S_{i2}(t|\mathbf{Z}_i, T_{i1}^m;\tau_1) - S_2(t;\tau_1)]. \qquad \text{(A.3)}
\end{aligned}
$$

The model for the second conditional gap time is fitted with subjects that have $I\{T_{i1} \le \tau_1 \wedge C_i\} = 1$. Denote $N_{\bullet 1}(\tau_1) = \sum_{i=1}^{n} I\{T_{i1} \le \tau_1 \wedge C_i\}$. By the results of Andersen and Gill (1982), we know that in (A.2),

$$
\begin{aligned}
& N_{\bullet 1}(\tau_1)^{1/2}[\widehat{S}_{i2}(t|\mathbf{Z}_i, T_{i1}^m;\tau_1) - S_{i2}(t;\tau_1)] \\
& \quad = - S_{i2}(t|\mathbf{W}_i^m;\tau_1) \exp(\boldsymbol{\theta}_2' \mathbf{W}_i^m) \{ \int_0^t [\mathbf{W}_i^m - \overline{\mathbf{z}}_2(u, \boldsymbol{\theta}_2)] d\Lambda_{02}(u;\tau_1) \}' \boldsymbol{\Sigma}_2(\boldsymbol{\theta})^{-1} \\
& \qquad \times N_{\bullet 1}(\tau_1)^{-1/2} \sum_{k=1}^{n} N_{k1}(\tau_1) \int_0^{\tau_2} [\mathbf{W}_k^m - \overline{\boldsymbol{z}}_2(u, \boldsymbol{\theta}_2)] d\widetilde{M}_{k2}(u) \\
& \qquad - S_{i2}(t|\mathbf{W}_i^m;\tau_1) \exp(\boldsymbol{\theta}_2' \mathbf{W}_i^m) N_{\bullet 1}(\tau_1)^{-1/2} \\
& \qquad \times \sum_{k=1}^{n} N_{k1}(\tau_1) \int_0^t S_2^{(0)}(u, \boldsymbol{\theta}_2)^{-1} d\widetilde{M}_{k2}(u) + o_p(1) \\
& \quad = N_{\bullet 1}(\tau_1)^{-1/2} \sum_{k=1}^{n} N_{k1}(\tau_1) \gamma_{ik1}^m(t) + o_p(1), \qquad \text{(A.4)}
\end{aligned}
$$

where

$$
\begin{aligned}
\gamma_{ik1}^m(t) = & - S_{i2}(t|\mathbf{W}_i^m;\tau_1) \exp(\boldsymbol{\theta}_2' \mathbf{W}_i^m) \{ \int_0^t [\mathbf{W}_i^m - \overline{\mathbf{z}}_2(u, \boldsymbol{\theta}_2)] d\Lambda_{02}(u;\tau_1) \}' \boldsymbol{\Sigma}_2(\boldsymbol{\theta})^{-1} \\
& \times \int_0^{\tau_2} [\mathbf{W}_k^m - \overline{\boldsymbol{z}}_2(u, \boldsymbol{\theta}_2)] d\widetilde{M}_{k2}(u) \\
& - S_{i2}(t|\mathbf{W}_i^m;\tau_1) \exp(\boldsymbol{\theta}_2' \mathbf{W}_i^m) \int_0^t S_2^{(0)}(u, \boldsymbol{\theta}_2)^{-1} d\widetilde{M}_{k2}(u).
\end{aligned}
$$

Plugging (A.4) into (A.2) then changing the order of summation, we have

$$\begin{aligned} n^{1/2}M^{-1}\sum_{m=1}^{M}N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{i=1}^{n}N_{i1}^{m}(\tau_1)[\widehat{S}_{i2}(t|\mathbf{Z}_i,T_{i1}^{m};\tau_1)-S_{i2}(t|\mathbf{Z}_i,T_{i1}^{m};\tau_1)] \\ =[n/N_{\bullet 1}(\tau_1)]^{1/2}M^{-1}\sum_{m=1}^{M}N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{i=1}^{n}N_{i1}^{m}(\tau_1)N_{\bullet 1}(\tau_1)^{-1/2} \\ \times\sum_{k=1}^{n}N_{k1}(\tau_1)\gamma_{ik1}^{m}(t)+o_p(1) \\ =n^{-1/2}\sum_{i=1}^{n}M^{-1}\sum_{m=1}^{M}P(T_{i1}\le\tau_1\wedge C_i)^{-1}N_{i1}(\tau_1)\phi_{i1}^{m}(t)+o_p(1), \end{aligned} \tag{A.5}$$

where

$$\phi_{i1}^{m}(t)=N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{k=1}^{n}N_{k1}^{m}(\tau_1)\gamma_{ki1}^{m}(t).$$

A more compact way to write $\phi_{i1}^{m}(t)$ is given by,

$$\begin{aligned} \phi_{i1}^{m}(t)=&\mathbf{b}_m(t)'\int_0^{\tau_2}[\mathbf{W}_i^{m}-\overline{\boldsymbol{z}}_2(u)]d\widetilde{M}_{i2}(u)-E[S_2(t|\mathbf{W}_i)\exp(\boldsymbol{\theta}_2'\mathbf{W}_i)|T_1^{m}\le\tau_1] \\ &\times\int_0^{t}\frac{1}{s_2^{(0)}(u,\boldsymbol{\theta}_2)}d\widetilde{M}_{i2}(u), \end{aligned}$$

where

$$\begin{aligned} \mathbf{b}_m(t)=&\boldsymbol{\Sigma}_2(\boldsymbol{\theta})^{-1}\int_0^{t}\{\overline{\boldsymbol{z}}_2(u,\boldsymbol{\theta}_2)E[S_2(t|\mathbf{W}_i)\exp(\boldsymbol{\theta}_2'\mathbf{W}_i)|T_1^{m}\le\tau_1] \\ &-E[\mathbf{W}_iS_2(t|\mathbf{W}_i)\exp(\boldsymbol{\theta}_2'\mathbf{W}_i)|T_1^{m}\le\tau_1]\}d\Lambda_{02}(u;\tau_1). \end{aligned}$$

We can then write,

$$\begin{aligned} &n^{1/2}M^{-1}\sum_{m=1}^{M}N_{\bullet 1}^{m}(\tau_1)^{-1}\sum_{i=1}^{n}N_{i1}^{m}(\tau_1)[S_{i2}(t|\mathbf{Z}_i,T_{i1}^{m};\tau_1)-S_2(t;\tau_1)] \\ =&n^{-1/2}\sum_{i=1}^{n}M^{-1}\sum_{m=1}^{M}P(T_{i1}\le\tau_1)^{-1}N_{i1}^{m}(\tau_1)[S_{i2}(t|\mathbf{W}_i^{m};\tau_1)-S_2(t;\tau_1)]+o_p(1). \end{aligned} \tag{A.6}$$

Thus, combining (A.5) and (A.6), $n^{1/2}[\widehat{S}_2(t;\tau_1)-S_2(t;\tau_1)]$ has the following linear representation as indicated in Theorem 1.1,

$$n^{1/2}[\widehat{S}_2(t;\tau_1)-S_2(t;\tau_1)]=n^{-1/2}\sum_{i=1}^{n}\varphi_{i1}(t)+o_p(1), \tag{A.7}$$

where

$$\varphi_{i1}(t) = M^{-1}\sum_{m=1}^{M}\{P(T_{i1} \le \tau_1 \wedge C_i)^{-1}N_{i1}(\tau_1)\phi_{i1}^m(t) + P(T_{i1} \le \tau_1)^{-1}N_{i1}^m(\tau_1)[S_{i2}(t|\mathbf{W}_i^m;\tau_1) - S_2(t;\tau_1)]\} + o_p(1).$$

By integrating both sides of (A.7), $n^{1/2}[\widehat{\mu}_2(L;\tau_1) - \mu_2(L;\tau_1)] = \int_0^L[\widehat{S}_2(t;\tau_1) - S_2(t;\tau_1)]dt$ has the following linear representation,

$$n^{1/2}[\widehat{\mu}_2(L;\tau_1) - \mu_2(L;\tau_1)] = n^{-1/2}\sum_{i=1}^{n}\varphi_{i2}(L) + o_p(1),$$

where

$$\varphi_{i2}(L) = \int_0^L \varphi_{i1}(t)dt.$$

A.3.6 Proofs of Theorems 1.1, 1.2 and 1.3

The variates $\varphi_{i1}(t)$ are independent and identically distributed mean-zero random variables (for $(i = 1, \cdots, n)$, for any fixed t), such that $E[\varphi_{i1}(t)^2] < \infty$; the same holds for $\varphi_{i2}(L)$, such that $E[\varphi_{i2}(L)^2] < \infty$. Thus, $n^{1/2}[\widehat{S}_2(t;\tau_1) - S_2(t;\tau_1)]$ and $n^{1/2}[\widehat{\mu}_2(L;\tau_1) - \mu_2(L;\tau_1)]$ are asymptotically normal with means 0 and variances $E[\varphi_{i1}(t)^2]$ and $E[\varphi_{i2}(L)^2]$, respectively.

Hence, Theorem 1.1 is proved. The proof of Theorem 1.2 is very similar and is therefore not shown here. Theorem 1.3 follows naturally by combining the results from Theorems 1.1 and 1.2.

A.4 Additional Simulations

In addition to the simulation studies in Section 1.4, two more data configurations are considered. We aim to generate gap times from more similar Cox models, so the true differences between gap times are less considerable. Subjects are $i = 1, \cdots, n$, and each subject has two independent binary covariates, Z_{i1} and Z_{i2}, with $\Pr\{Z_{i1} =$

$1\} = \Pr\{Z_{i2} = 1\} = 0.5$. Two gap times T_{i1} and $\widetilde{T}_{i2}$ are generated from the following proportional hazards models:

$$\lambda_{i1}(t) = \lambda_{01}(t) \exp\{\beta_1 Z_{i1} + \beta_2 Z_{i2}\}$$
$$\lambda_{i2}(t) = \lambda_{02}(t) \exp\{\beta_3 Z_{i1} + \beta_4 Z_{i2} + \beta_5 T_{i1}\},$$

where $\lambda_{01}(t) = \lambda_{02}(t) = 0.6$, $\beta_1 = \beta_3 = \log(1.5)$, and $\beta_2 = \beta_4 = -\log(1.5)$.

The censoring time, C_i, follows a uniform distribution on $(0, 12)$. We set $L = \tau_1 = 5$, and the number of multiple imputations to $M = 5$. For each data configuration, the sample size was $n = 250$ and we ran 1,000 replicates. We consider three cases, which are different based on the choices of β_5; values used included $\beta_5 = 0$, $\beta_5 = \log(1.025)$, and $\beta_5 = -\log(1.025)$. Table A.1 lists the simulation results for each of the three cases. Table A.2 shows the results where a larger correlation exists between the two gap times, by setting $\beta_1 = \beta_3 = \log(2.5)$ and $\beta_2 = \beta_4 = -\log(2.5)$. Three cases are considered under this scenario again.

In the above additional simulation studies, we observe similar trends as in Table 1.1 in Section 1.4 of the main chapter.

We carried out two more sets of simulations that incorporate $\log(T_{i1} + 1)$ as a covariate in the Cox model for the second gap time. In the first set of simulations, the model is correctly specified when estimating the survival functions and the restricted mean lifetimes (see Table A.3); while for the second set, we still used the linear term of T_{i1} in the Cox model, so that there is model mis-specification (results shown in Table A.4). Again, the biases are quite small in both sets. The ESDs and ASEs agree with each other, and the coverage probabilities are close to 0.95. Comparing results in Table A.3 with those in Table A.4, we see very little differences, indicating that the model mis-specification is not a deal breaker in the data configuration we

Table A.1: Additional simulation results for estimating survival functions and restricted mean lifetimes: more similar Cox models

Setting*	Parameter	True	BIAS	ESD	ASE	ECP
1	$\mu_1(L;\tau_1)$	1.59	-0.003	0.090	0.093	0.95
	$\mu_2(L;\tau_1)$	1.59	-0.008	0.107	0.104	0.94
	$\Delta(L;\tau_1)$	0.00	-0.006	0.133	0.141	0.95
	$S_1(1;\tau_1)$	0.54	0.002	0.032	0.0320	0.94
	$S_2(1;\tau_1)$	0.54	0.001	0.035	0.035	0.95
	$\delta(1;\tau_1)$	0.00	-0.001	0.046	0.047	0.96
	$S_1(3;\tau_1)$	0.17	0.002	0.025	0.025	0.95
	$S_2(3;\tau_1)$	0.17	0.001	0.029	0.029	0.94
	$\delta(3;\tau_1)$	0.00	-0.001	0.036	0.0375	0.95
	$S_1(5;\tau_1)$	0.06	0.003	0.017	0.017	0.94
	$S_2(5;\tau_1)$	0.06	0.003	0.021	0.020	0.92
	$\delta(5;\tau_1)$	0.00	0.001	0.026	0.026	0.94
2	$\mu_1(L;\tau_1)$	1.59	-0.003	0.090	0.093	0.95
	$\mu_2(L;\tau_1)$	1.54	-0.007	0.105	0.102	0.94
	$\Delta(L;\tau_1)$	-0.05	-0.004	0.133	0.140	0.95
	$S_1(1;\tau_1)$	0.54	0.002	0.032	0.032	0.94
	$S_2(1;\tau_1)$	0.53	0.001	0.035	0.035	0.95
	$\delta(1;\tau_1)$	-0.01	-0.001	0.047	0.047	0.95
	$S_1(3;\tau_1)$	0.17	0.002	0.025	0.025	0.95
	$S_2(3;\tau_1)$	0.16	0.001	0.028	0.028	0.95
	$\delta(3;\tau_1)$	-0.01	-0.002	0.036	0.037	0.95
	$S_1(5;\tau_1)$	0.06	0.003	0.017	0.017	0.94
	$S_2(5;\tau_1)$	0.05	0.003	0.020	0.019	0.93
	$\delta(5;\tau_1)$	-0.01	0.000	0.026	0.025	0.94
3	$\mu_1(L;\tau_1)$	1.59	-0.003	0.090	0.093	0.95
	$\mu_2(L;\tau_1)$	1.63	-0.008	0.109	0.106	0.93
	$\Delta(L;\tau_1)$	0.05	-0.006	0.134	0.141	0.96
	$S_1(1;\tau_1)$	0.54	0.002	0.032	0.032	0.94
	$S_2(1;\tau_1)$	0.55	0.000	0.035	0.035	0.95
	$\delta(1;\tau_1)$	0.01	-0.002	0.046	0.047	0.96
	$S_1(3;\tau_1)$	0.17	0.002	0.025	0.025	0.95
	$S_2(3;\tau_1)$	0.18	0.001	0.029	0.029	0.95
	$\delta(3;\tau_1)$	0.01	-0.001	0.036	0.038	0.95
	$S_1(5;\tau_1)$	0.06	0.003	0.017	0.017	0.94
	$S_2(5;\tau_1)$	0.07	0.003	0.022	0.020	0.92
	$\delta(5;\tau_1)$	0.01	0.000	0.028	0.026	0.94

* Settings 1, 2, and 3, correspond to $\beta_5 = 0$, $\beta_5 = \log(1.025)$, and $\beta_5 = -\log(1.025)$, respectively.

Table A.2: Additional simulation results for estimating survival functions and restricted mean lifetimes: stronger association between gap times

Setting*	Parameter	True	BIAS	ESD	ASE	ECP
1	$\mu_1(L;\tau_1)$	1.59	-0.011	0.089	0.098	0.95
	$\mu_2(L;\tau_1)$	1.59	0.030	0.115	0.112	0.94
	$\Delta(L;\tau_1)$	-0.00	0.041	0.131	0.154	0.97
	$S_1(1;\tau_1)$	0.51	0.000	0.031	0.032	0.95
	$S_2(1;\tau_1)$	0.50	0.010	0.036	0.035	0.93
	$\delta(1;\tau_1)$	-0.00	0.001	0.045	0.044	0.94
	$S_1(3;\tau_1)$	0.19	0.000	0.024	0.024	0.94
	$S_2(3;\tau_1)$	0.19	0.011	0.023	0.030	0.95
	$\delta(3;\tau_1)$	-0.00	0.011	0.035	0.036	0.94
	$S_1(5;\tau_1)$	0.08	0.002	0.017	0.017	0.95
	$S_2(5;\tau_1)$	0.09	0.008	0.024	0.023	0.94
	$\delta(5;\tau_1)$	0.00	0.006	0.028	0.028	0.94
2	$\mu_1(L;\tau_1)$	1.59	-0.011	0.089	0.010	0.95
	$\mu_2(L;\tau_1)$	1.55	-0.011	0.113	0.110	0.95
	$\Delta(L;\tau_1)$	-0.04	-0.004	0.131	0.154	0.98
	$S_1(1;\tau_1)$	0.51	0.000	0.031	0.032	0.95
	$S_2(1;\tau_1)$	0.49	0.000	0.036	0.035	0.94
	$\delta(1;\tau_1)$	-0.01	0.000	0.045	0.045	0.95
	$S_1(3;\tau_1)$	0.187	0.000	0.024	0.024	0.94
	$S_2(3;\tau_1)$	0.177	0.002	0.029	0.029	0.95
	$\delta(3;\tau_1)$	-0.01	0.002	0.035	0.036	0.94
	$S_1(5;\tau_1)$	0.08	0.002	0.017	0.017	0.95
	$S_2(5;\tau_1)$	0.08	0.002	0.023	0.022	0.93
	$\delta(5;\tau_1)$	-0.00	-0.001	0.028	0.027	0.93
3	$\mu_1(L;\tau_1)$	1.59	-0.011	0.089	0.098	0.95
	$\mu_2(L;\tau_1)$	1.63	-0.013	0.116	0.114	0.94
	$\Delta(L;\tau_1)$	0.04	-0.005	0.131	0.155	0.98
	$S_1(1;\tau_1)$	0.51	0.000	0.031	0.032	0.95
	$S_2(1;\tau_1)$	0.51	0.000	0.036	0.035	0.95
	$\delta(1;\tau_1)$	0.01	-0.000	0.044	0.044	0.95
	$S_1(3;\tau_1)$	0.19	0.000	0.024	0.024	0.94
	$S_2(3;\tau_1)$	0.10	0.001	0.030	0.030	0.95
	$\delta(3;\tau_1)$	0.01	0.001	0.035	0.036	0.94
	$S_1(5;\tau_1)$	0.08	0.002	0.017	0.017	0.95
	$S_2(5;\tau_1)$	0.09	0.002	0.025	0.024	0.93
	$\delta(5;\tau_1)$	0.01	-0.001	0.029	0.029	0.94

* $\beta_1 = \beta_3 = \log(2.5)$, $\beta_2 = \beta_4 = -\log(2.5)$. Settings 1, 2, and 3, correspond to $\beta_5 = 0$, $\beta_5 = \log(1.025)$, and $\beta_5 = -\log(1.025)$, respectively.

considered.

Overall, our method is demonstrated to work well under the scenarios considered.

A.5 Parameter Estimates for the First and Second Transplant Models

The parameter estimates for the fitted Cox models for the first and second gap times are listed in Tables A.5 and Table A.6, respectively. For example, the hazard of graft failure after the first transplant for subjects aged more than 60 are 43.3% larger than that for subjects aged between 18 and 40; while given the observed first gap time is less than 10 years, the hazard of graft failure after the second transplant for subjects aged more than 60 are 54.9% larger than that for subjects aged between 18 and 40. Covariate effects on $\lambda_{i1}(t)$ and $\lambda_{i2}(t)$ appear to be quite similar, in terms of magnitude and direction.

Table A.3: Additional simulation results for estimating survival functions and restricted mean lifetimes: correct model specification

	Setting 1					Setting 2				
Parameter	True	BIAS	ESD	ASE	ECP	True	BIAS	ESD	ASE	ECP
$\mu_1(L;\tau_1)$	3.094	-0.029	0.120	0.122	0.947	2.799	-0.009	0.123	0.128	0.948
$\mu_2(L;\tau_1)$	2.016	-0.006	0.156	0.149	0.939	1.848	-0.023	0.151	0.145	0.933
$\Delta(L;\tau_1)$	-1.078	0.023	0.189	0.195	0.950	-0.951	-0.014	0.175	0.200	0.974
$S_1(1;\tau_1)$	0.808	-0.002	0.026	0.026	0.950	0.754	0.002	0.030	0.030	0.948
$S_2(1;\tau_1)$	0.634	0.004	0.043	0.043	0.936	0.571	-0.001	0.044	0.043	0.945
$\delta(1;\tau_1)$	-0.173	0.005	0.049	0.050	0.944	-0.183	-0.003	0.050	0.051	0.947
$S_1(3;\tau_1)$	0.532	-0.005	0.034	0.034	0.950	0.460	0.001	0.034	0.034	0.943
$S_2(3;\tau_1)$	0.270	0.002	0.044	0.042	0.934	0.240	-0.001	0.039	0.039	0.947
$\delta(3;\tau_1)$	-0.263	0.007	0.053	0.054	0.948	-0.220	-0.001	0.048	0.049	0.952
$S_1(5;\tau_1)$	0.352	-0.001	0.034	0.034	0.952	0.302	0.002	0.030	0.030	0.956
$S_2(5;\tau_1)$	0.121	0.005	0.035	0.034	0.929	0.122	0.003	0.033	0.032	0.940
$\delta(5;\tau_1)$	-0.232	0.006	0.047	0.048	0.954	-0.180	0.001	0.042	0.042	0.951

	Setting 3					Setting 4				
Parameter	True	BIAS	ESD	ASE	ECP	True	BIAS	ESD	ASE	ECP
$\mu_1(L;\tau_1)$	2.146	-0.021	0.106	0.109	0.953	2.023	-0.005	0.101	0.110	0.960
$\mu_2(L;\tau_1)$	3.177	-0.028	0.146	0.141	0.927	2.967	-0.010	0.151	0.143	0.939
$\Delta(L;\tau_1)$	1.030	-0.005	0.170	0.179	0.960	0.946	-0.006	0.164	0.186	0.973
$S_1(1;\tau_1)$	0.659	0.000	0.030	0.031	0.954	0.608	0.002	0.033	0.032	0.934
$S_2(1;\tau_1)$	0.818	-0.001	0.030	0.029	0.944	0.779	0.001	0.032	0.031	0.944
$\delta(1;\tau_1)$	0.159	-0.001	0.040	0.042	0.960	0.172	-0.003	0.041	0.042	0.953
$S_1(3;\tau_1)$	0.300	-0.004	0.030	0.030	0.949	0.278	0.002	0.027	0.028	0.954
$S_2(3;\tau_1)$	0.553	-0.002	0.039	0.040	0.948	0.502	0.002	0.041	0.040	0.939
$\delta(3;\tau_1)$	0.253	0.002	0.047	0.050	0.960	0.224	-0.001	0.046	0.046	0.951
$S_1(5;\tau_1)$	0.141	0.000	0.024	0.024	0.946	0.149	0.002	0.021	0.022	0.948
$S_2(5;\tau_1)$	0.376	0.001	0.044	0.042	0.941	0.345	0.002	0.042	0.041	0.941
$\delta(5;\tau_1)$	0.235	0.001	0.049	0.049	0.941	0.196	0.001	0.045	0.046	0.948

Setting 1: $\lambda_{01}(t)$=0.2, $\lambda_{02}(t)$=0.4, $\beta_1 = \beta_3 = \log(1.5)$, $\beta_2 = \beta_4 = -\log(1.5)$, $\beta_5 = \log(1.05)$
Setting 2: $\lambda_{01}(t)$=0.2, $\lambda_{02}(t)$=0.4, $\beta_1 = \beta_3 = \log(2.5)$, $\beta_2 = \beta_4 = -\log(2.5)$, $\beta_5 = \log(1.05)$
Setting 3: $\lambda_{01}(t)$=0.4, $\lambda_{02}(t)$=0.2, $\beta_1 = \beta_3 = \log(1.5)$, $\beta_2 = \beta_4 = -\log(1.5)$, $\beta_5 = -\log(1.05)$
Setting 4: $\lambda_{01}(t)$=0.4, $\lambda_{02}(t)$=0.2, $\beta_1 = \beta_3 = \log(2.5)$, $\beta_2 = \beta_4 = -\log(2.5)$, $\beta_5 = -\log(1.05)$

Table A.4: Additional simulation results for estimating survival functions and restricted mean lifetimes: incorrect model specification

	Setting 1					Setting 2				
Parameter	True	BIAS	ESD	ASE	ECP	True	BIAS	ESD	ASE	ECP
$\mu_1(L;\tau_1)$	3.094	-0.029	0.120	0.122	0.947	2.799	-0.009	0.123	0.128	0.948
$\mu_2(L;\tau_1)$	2.016	-0.006	0.156	0.149	0.940	1.848	-0.023	0.152	0.145	0.932
$\Delta(L;\tau_1)$	-1.078	0.023	0.189	0.195	0.953	-0.951	-0.014	0.175	0.200	0.971
$S_1(1;\tau_1)$	0.808	-0.002	0.026	0.026	0.949	0.754	0.002	0.030	0.030	0.948
$S_2(1;\tau_1)$	0.634	0.004	0.043	0.043	0.936	0.571	-0.001	0.044	0.043	0.945
$\delta(1;\tau_1)$	-0.173	0.005	0.049	0.050	0.946	-0.183	-0.003	0.051	0.051	0.948
$S_1(3;\tau_1)$	0.532	-0.004	0.034	0.034	0.950	0.460	0.001	0.034	0.034	0.943
$S_2(3;\tau_1)$	0.270	0.002	0.044	0.042	0.936	0.240	-0.001	0.039	0.039	0.945
$\delta(3;\tau_1)$	-0.263	0.007	0.053	0.054	0.950	-0.220	-0.002	0.048	0.049	0.952
$S_1(5;\tau_1)$	0.352	-0.001	0.034	0.034	0.952	0.302	0.002	0.030	0.030	0.956
$S_2(5;\tau_1)$	0.121	0.005	0.035	0.034	0.931	0.122	0.003	0.034	0.032	0.938
$\delta(5;\tau_1)$	-0.232	0.007	0.047	0.048	0.954	-0.180	0.001	0.042	0.042	0.950

	Setting 3					Setting 4				
Parameter	True	BIAS	ESD	ASE	ECP	True	BIAS	ESD	ASE	ECP
$\mu_1(L;\tau_1)$	2.146	-0.021	0.106	0.109	0.953	2.023	-0.005	0.101	0.110	0.960
$\mu_2(L;\tau_1)$	3.177	-0.027	0.146	0.141	0.927	2.967	-0.009	0.151	0.143	0.939
$\Delta(L;\tau_1)$	1.030	-0.005	0.170	0.180	0.959	0.946	-0.006	0.164	0.186	0.975
$S_1(1;\tau_1)$	0.659	0.000	0.030	0.031	0.954	0.608	0.002	0.033	0.032	0.934
$S_2(1;\tau_1)$	0.818	-0.001	0.030	0.029	0.944	0.779	0.001	0.032	0.031	0.943
$\delta(1;\tau_1)$	0.159	-0.001	0.040	0.042	0.961	0.172	-0.003	0.041	0.042	0.954
$S_1(3;\tau_1)$	0.300	-0.004	0.030	0.030	0.949	0.278	0.002	0.027	0.028	0.954
$S_2(3;\tau_1)$	0.553	-0.002	0.039	0.040	0.946	0.502	0.002	0.041	0.040	0.939
$\delta(3;\tau_1)$	0.253	0.002	0.047	0.050	0.959	0.224	-0.000	0.046	0.046	0.951
$S_1(5;\tau_1)$	0.141	0.000	0.024	0.024	0.946	0.149	0.002	0.021	0.022	0.948
$S_2(5;\tau_1)$	0.376	0.001	0.044	0.042	0.938	0.345	0.003	0.042	0.041	0.939
$\delta(5;\tau_1)$	0.235	0.001	0.049	0.049	0.938	0.196	0.001	0.046	0.046	0.949

Setting 1: $\lambda_{01}(t)$=0.2, $\lambda_{02}(t)$=0.4, $\beta_1 = \beta_3 = \log(1.5)$, $\beta_2 = \beta_4 = -\log(1.5)$, $\beta_5 = \log(1.05)$
Setting 2: $\lambda_{01}(t)$=0.2, $\lambda_{02}(t)$=0.4, $\beta_1 = \beta_3 = \log(2.5)$, $\beta_2 = \beta_4 = -\log(2.5)$, $\beta_5 = \log(1.05)$
Setting 3: $\lambda_{01}(t)$=0.4, $\lambda_{02}(t)$=0.2, $\beta_1 = \beta_3 = \log(1.5)$, $\beta_2 = \beta_4 = -\log(1.5)$, $\beta_5 = -\log(1.05)$
Setting 4: $\lambda_{01}(t)$=0.4, $\lambda_{02}(t)$=0.2, $\beta_1 = \beta_3 = \log(2.5)$, $\beta_2 = \beta_4 = -\log(2.5)$, $\beta_5 = -\log(1.05)$

Table A.5: Analysis of SRTR data: Parameter estimates for the first-kidney-transplant Cox model

Covariate		$\widehat{\beta}$	SE($\widehat{\beta}$)	$\widehat{HR}$	P-value
Gender	Male	0.134	0.011	1.144	< 0.0001
Race	White	0	0	1	-
	Asian	-0.359	0.027	0.699	< 0.0001
	African American	0.262	0.012	1.299	< 0.0001
	Hispanic or Latino	-0.158	0.017	0.854	< 0.0001
	Other	-0.037	0.039	0.964	0.339
Diabetes	No	0	0	1	-
	Insulin Dependent	0.354	0.014	1.425	< 0.0001
	Non-Insulin Dependent	0.322	0.018	1.380	< 0.0001
BMI	[25,30)	0	0	1	-
	[0,20)	0.078	0.022	1.081	0.0005
	[20,25)	-0.005	0.013	0.995	0.721
	[30,35)	0.076	0.014	1.079	< 0.0001
	[35,60]	0.190	0.018	1.210	< 0.0001
Waiting Time	$[0, 1)$ Year	0	0	1	-
	$[1, 2)$ Year	-0.034	0.014	0.966	0.012
	$[2, 4)$ Year	-0.026	0.013	0.974	0.041
	$>= 4$ Year	0.060	0.016	1.062	0.0002
Age	$[18, 40)$	0	0	1	-
	$[40, 50)$	-0.086	0.017	0.918	< 0.0001
	$[50, 60)$	0.063	0.016	1.065	< 0.0001
	$>= 60$	0.360	0.016	1.433	< 0.0001
Donor Died from Stroke	No	0	0	1	-
	Yes	0.100	0.012	1.105	< 0.0001
Donor Hypertension	No	0	0	1	-
	Yes	0.138	0.013	1.148	< 0.0001
Donor Diabetes	No	0	0	1	-
	Yes, 0-5 years	0.156	0.028	1.169	< 0.0001
	Yes, 6-10 years	0.372	0.044	1.450	< 0.0001
	Yes, > 10 years	0.356	0.047	1.428	< 0.0001
	Yes, duration unknown	0.140	0.055	1.150	0.012
Donor BMI	[20,25)	0	0	1	-
	$[0, 20)$	0.076	0.017	1.079	< 0.0001
	$[25, 30)$	-0.020	0.013	0.980	0.109
	$[30, 35)$	-0.016	0.016	0.984	0.326
	$[35, 60]$	0.048	0.019	1.049	0.012
Donor Creatinine	$[0, 1)$	0	0	1	-
	$[1, 2)$	0.049	0.011	1.051	< 0.0001
	$>= 2$	0.080	0.022	1.084	0.0003
Donor Age*		0.019	0.001	1.019	< 0.0001
Calendar Year		-0.034	0.002	0.966	< 0.0001
PRA	$= 0$	0	0	1	-
	$(0, 20]$	0.059	0.015	1.060	< 0.0001
	$(20, 100]$	0.166	0.017	1.181	< 0.0001

* Donor Age $= (DA - 40) * I(DA > 40)$, where DA represents original donor age.

Table A.6: Analysis of SRTR data: Parameter estimates for the re-kidney-transplant Cox model

Covariate		$\widehat{\beta}$	SE($\widehat{\beta}$)	$\widehat{HR}$	P-value
Gender	Male	0.104	0.075	1.109	0.169
Race	White	0	0	1	-
	Asian	-0.479	0.233	0.619	0.040
	African American	0.303	0.079	1.354	0.0001
	Hispanic or Latino	-0.400	0.164	0.670	0.015
	Other	0.349	0.324	1.417	0.281
Diabetes	No	0	0	1	-
	Insulin Dependent	0.306	0.127	1.358	0.016
	Non-Insulin Dependent	-0.290	0.224	0.749	0.196
BMI	[25,30)	0	0	1	-
	[0,20)	-0.123	0.180	0.884	0.492
	[20,25)	0.020	0.092	1.020	0.829
	[30,35)	0.243	0.098	1.275	0.013
	[35,60]	0.282	0.128	1.325	0.027
T_1^*		-0.194	0.062	0.824	0.002
Waiting Time	$[0, 1)$ Year	0	0	1	-
	$[1, 2)$ Year	0.013	0.100	1.013	0.896
	$[2, 4)$ Year	0.094	0.098	1.099	0.333
	$>= 4$ Year	0.272	0.120	1.313	0.024
Age	$[18, 40)$	0	0	1	-
	$[40, 50)$	0.017	0.107	1.017	0.875
	$[50, 60)$	0.112	0.103	1.119	0.275
	$>= 60$	0.437	0.108	1.549	< 0.0001
Donor Died from Stroke	No	0	0	1	-
	Yes	0.124	0.082	1.132	0.130
Donor Hypertension	No	0	0	1	-
	Yes	0.315	0.092	1.370	0.0006
Donor Diabetes	No	0	0	1	-
	Yes, 0-5 years	-0.039	0.183	0.962	0.832
	Yes, 6-10 years	0.303	0.342	1.355	0.375
	Yes, $>$ 10 years	0.408	0.306	1.504	0.182
	Yes, duration unknown	0.241	0.387	1.273	0.533
Donor BMI	[20,25)	0	0	1	-
	[0,20)	-0.053	0.127	0.948	0.676
	[25,30)	-0.226	0.091	0.797	0.013
	[30,35)	-0.106	0.115	0.899	0.355
	[35,60]	0.181	0.131	1.198	0.168
Donor Creatinine	$[0, 1)$	0	0	1	-
	$[1, 2)$	0.024	0.076	1.024	0.752
	$>= 2$	0.394	0.161	1.483	0.014
Donor Age*		0.025	0.005	1.025	< 0.0001
Calendar Year		-0.037	0.015	0.963	0.014
PRA	$= 0$	0	0	1	-
	$(0, 20]$	-0.155	0.113	0.856	0.169
	$(20, 100]$	0.101	0.083	1.106	0.227

* Donor Age $= (DA - 40) * I(DA > 40)$, where DA represents the original donor age; $T_1^* = (T_1 - 4) * I(T_1 > 4)$ (in years).

APPENDIX B

Supplementary Materials for Chapter II

In this appendix, we prove Theorem 2.1. We first give a review of the notation.

B.1 Notation

i = subject $(i = 1, \cdots, n)$

T_{i1} = first gap time

$\widetilde{T}_{i2}$ = second gap time

C_i = censoring time for T_{i1}

$\widetilde{C}_{i2} = C_i - T_{i1}$ = censoring time for $\widetilde{T}_{i2}$

$\tau = \sup\{t : P(C_i \geq t) > 0\}$

$t_1 = \tau/2$

$\mathbf{Z}_i$ = time-constant covariate vector

$\lambda_{i1}(t) = \lim_{\delta \to 0} \delta^{-1} P(t < T_{i1} \leq t + \delta | T_{i1} > t)$

$\lambda_{i1}(t) = \lambda_0(t) \exp\{\boldsymbol{\beta}' \mathbf{Z}_i\}$

$\lambda_{i2}(t; t_1) = \lim_{\delta \to 0} \delta^{-1} P(t < \widetilde{T}_{i2} \leq t + \delta | \widetilde{T}_{i2} \geq t, T_{i1} \leq t_1)$

$\lambda_{i2}(t; t_1) = \lambda_{i1}(t) e^{\theta}$

$\Lambda_{i1}(t) = \int_0^t \lambda_{i1}(ds)$

$\Lambda_{i2}(t; t_1) = \int_0^t \lambda_{i2}(ds; t_1)$

$\lambda_i^C(t) = \lambda_0^C(t) \exp\{\boldsymbol{\alpha}' \mathbf{Z}_i^C\}$

$N_{i1}(t) = I\{T_{i1} \leq t \wedge C_i\}$

$Y_{i1}(t) = I\{T_{i1} \wedge C_i \geq t\}$

$N_{i2}(t;t_1) = I\{\widetilde{T}_{i2} \leq t \wedge \widetilde{C}_{i2}, T_{i1} \leq t_1\}$

$Y_{i2}(t;t_1) = I\{\widetilde{T}_{i2} \wedge \widetilde{C}_{i2} \geq t, T_{i1} \leq t_1\}$

$N_i^C(t) = I\{C_i \leq T_{i1} + \widetilde{T}_{i2}, C_i \leq t\}$

$Y_i^C(t) = I\{C_i \wedge (T_{i1} + \widetilde{T}_{i2}) \geq t\}$

$M_{i1}(t) = N_{i1}(t) - \int_0^t \lambda_{i1}(u) Y_{i1}(u) du$

$M_{i2}(t;t_1) = N_{i2}(t;t_1) - \int_0^t \lambda_{i2}(u;\tau_1) Y_{i2}(u;t_1) du$

$M_i^C(t) = N_i^C(t) - \int_0^t \lambda_i^C(u) Y_i^C(u) du$

$W_{i2}(t;t_1) = Y_{i2}(t;t_1) P(C_i \geq t + T_{i1} | T_{i1})^{-1}$

$s_C^{(d)}(t,\boldsymbol{\alpha}) = E[Y_i^C(t) \boldsymbol{Z}_i^{C\otimes d} \exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}], d = 0,1,2$

$s_1^{(d)}(t,\boldsymbol{\beta}) = E[Y_{i1}(t) \boldsymbol{Z}_i^{\otimes d} \exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}], d = 0,1,2$

$S_C^{(d)}(t,\boldsymbol{\alpha}) = n^{-1} \sum_{i=1}^n [Y_i^C(t) \boldsymbol{Z}_i^{C\otimes d} \exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}], d = 0,1,2$

$S_1^{(d)}(t,\boldsymbol{\beta}) = n^{-1} \sum_{i=1}^n [Y_{i1}(t) \boldsymbol{Z}_i^{\otimes d} \exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}], d = 0,1,2$

$\overline{\boldsymbol{z}}_C(t,\boldsymbol{\alpha}) = \boldsymbol{s}_C^{(1)}(t,\boldsymbol{\alpha}) / s_C^{(0)}(t,\boldsymbol{\alpha})$

$\overline{\boldsymbol{z}}_1(t,\boldsymbol{\alpha}) = \boldsymbol{s}_1^{(1)}(t,\boldsymbol{\alpha}) / s_1^{(0)}(t,\boldsymbol{\alpha})$

$\widehat{A}_n(t;\widehat{W}_{i2}) = n^{-1} \sum_{i=1}^n \int_0^t \widehat{W}_{i2}(s;t_1) N_{i2}(ds;t_1)$

$\widehat{B}_n(t;\widehat{W}_{i2},\widehat{\Lambda}_{i1}) = n^{-1} \sum_{i=1}^n \int_0^t \widehat{W}_{i2}(s;t_1) \widehat{\Lambda}_{i1}(ds)$

$\widehat{\theta} = \log\left\{\widehat{A}_n(t;\widehat{W}_{i2}) / \widehat{B}_n(t;\widehat{W}_{i2},\widehat{\Lambda}_{i1})\right\}$

Next, we list the assumed conditions underlying our proofs.

B.2 Regularity Conditions

We assume that the following regularity conditions hold for $i = 1, \cdots, n$, $0 \leq s \leq t_1$, $0 \leq u \leq t_1$:

(a) $\{N_{i1}(\cdot), N_{i2}(\cdot), Y_{i1}(\cdot), Y_{i2}(\cdot), \mathbf{Z}_i\}$ are independent and identically distributed;

(b) $E[Y_{i1}(s)] > 0$ and $E[Y_{i2}(u; t_1)] > 0$;

(c) elements of $\mathbf{Z}_i$ are bounded almost surely;

(d) $\Lambda_{01}(s) < \infty$ and $\Lambda_{02}(u; t_1) < \infty$;

(e) positive-definiteness of the following matrices:

$$\begin{aligned}\boldsymbol{\Sigma}_C(\boldsymbol{\alpha}) =& E\left[\int_0^\tau \left\{\frac{\boldsymbol{s}_C^{(2)}(t,\boldsymbol{\alpha})}{s_C^{(0)}(t,\boldsymbol{\alpha})} - \overline{\boldsymbol{z}}_C(t,\boldsymbol{\alpha})^{\otimes 2}\right\} dN_i^C(t)\right],\\ \boldsymbol{\Sigma}_1(\boldsymbol{\beta}) =& E\left[\int_0^\tau \left\{\frac{\boldsymbol{s}_1^{(2)}(t,\boldsymbol{\beta})}{s_1^{(0)}(t,\boldsymbol{\beta})} - \overline{\boldsymbol{z}}_1(t,\boldsymbol{\beta})^{\otimes 2}\right\} dN_{i1}(t)\right].\end{aligned}$$

B.3 Proof of Theorem 2.1

The derivation is composed of several key parts. We work with each park separately and combine the results sequentially.

B.3.1 $n^{1/2}(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})$

Following the standard asymptotic results of the Cox model (Anderson and Gill 1982; Fleming and Harrington 1991; Andersen et al. 1993), we have,

$$n^{1/2}(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) = \boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1} n^{-1/2} \sum_{i=1}^n \int_0^\tau \{\boldsymbol{Z}_i^C - \overline{\boldsymbol{z}}_C(t;\boldsymbol{\alpha})\} dM_i^C(t;\boldsymbol{\alpha}) + o_p(1).$$

B.3.2 $n^{1/2}[\widehat{\Lambda}_i^C(t;\widehat{\boldsymbol{\alpha}}) - \Lambda_i^C(t)]$

We can decompose the quantity as following,

$$\begin{aligned} & n^{1/2}[\widehat{\Lambda}_i^C(t;\widehat{\boldsymbol{\alpha}}) - \Lambda_i^C(t;\boldsymbol{\alpha})] \\ =& n^{1/2}[\widehat{\Lambda}_i^C(t;\widehat{\boldsymbol{\alpha}}) - \widehat{\Lambda}_i^C(t;\boldsymbol{\alpha})] \qquad \text{(B.1)} \\ &+ n^{1/2}[\widehat{\Lambda}_i^C(t;\boldsymbol{\alpha}) - \Lambda_i^C(t)]. \qquad \text{(B.2)}\end{aligned}$$

The first term, (B.1), can be expressed as

$$
\begin{aligned}
&n^{1/2}[\widehat{\Lambda}_i^C(t;\widehat{\boldsymbol{\alpha}}) - \widehat{\Lambda}_i^C(t;\boldsymbol{\alpha})] \\
&=n^{-1/2}\sum_{i=1}^{n}\int_0^t \left[\exp\{\widehat{\boldsymbol{\alpha}}'\mathbf{Z}_i^C\}S_C^{(0)}(s;\widehat{\boldsymbol{\alpha}})^{-1} - \exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}S_C^{(0)}(s;\boldsymbol{\alpha})^{-1}\right] dN_i^C(s) \\
&=\boldsymbol{k}_{iC}'(t;\boldsymbol{\alpha})n^{1/2}(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}) + o_p(1) \\
&=\boldsymbol{k}_{iC}'(t;\boldsymbol{\alpha})\boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1}n^{-1/2}\sum_{i=1}^{n}\int_0^\tau \{\boldsymbol{Z}_i^C - \overline{\boldsymbol{z}}(t;\boldsymbol{\alpha})\}dM_i^C(t;\boldsymbol{\alpha}),
\end{aligned}
$$

where

$$
\boldsymbol{k}_{iC}(t;\boldsymbol{\alpha}) = \exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}\int_0^t \left\{\mathbf{Z}_i^C - \overline{\mathbf{z}}_C(t;\boldsymbol{\alpha})\right\} d\Lambda_0^C(t).
$$

The second term, (B.2), can be written as

$$
\begin{aligned}
&n^{1/2}[\widehat{\Lambda}_i^C(t;\boldsymbol{\alpha}) - \Lambda_i^C(t)] \\
&\quad= \exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}n^{1/2}\{\widehat{\Lambda}_0^C(t;\boldsymbol{\alpha}) - \Lambda_0^C(t)\} \\
&\quad= \exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}n^{-1/2}\sum_{i=1}^{n}\int_0^t s_C^{(0)}(s;\boldsymbol{\alpha})^{-1}dM_i^C(s;\boldsymbol{\alpha}) + o_p(1).
\end{aligned}
$$

Combining the results, we have

$$
\begin{aligned}
n^{1/2}[\widehat{\Lambda}_i^C(t;\widehat{\boldsymbol{\alpha}}) - \Lambda_i^C(t)] =&\boldsymbol{k}_{iC}'(t;\boldsymbol{\alpha})\boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1}n^{-1/2}\sum_{i=1}^{n}\int_0^\tau \{\boldsymbol{Z}_i^C - \overline{\boldsymbol{z}}(t;\boldsymbol{\alpha})\}dM_i^C(t;\boldsymbol{\alpha}) \\
&+ \exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}n^{-1/2}\sum_{i=1}^{n}\int_0^t s_C^{(0)}(s;\boldsymbol{\alpha})^{-1}dM_i^C(s;\boldsymbol{\alpha}) + o_p(1).
\end{aligned}
$$

B.3.3 $n^{1/2}[\widehat{W}_{i2}(t;\widehat{\Lambda}_i^C) - W_{i2}(t)]$

Because $W_{i2}(t) = Y_{i2}(t;t_1)P(C_i \geq t+T_{i1}|\mathbf{Z}_i, T_{i1})^{-1} = Y_{i2}(t;t_1)\exp\{\Lambda_i^C(t+T_{i1}|T_{i1})\}$, we have

$$
n^{1/2}[\widehat{W}_{i2}(t;\widehat{\Lambda}_i^C) - W_{i2}(t)] = n^{1/2}Y_{i2}(t;t_1)\left[\exp\{\widehat{\Lambda}_i^C(t+T_{i1}|T_{i1})\} - \exp\{\Lambda_i^C(t+T_{i1}|T_{i1})\}\right]
$$

By a Taylor series expansion and previous results,

$$
\begin{aligned}
& n^{1/2}[\widehat{W}_{i2}(t;\widehat{\Lambda}_i^C) - W_{i2}(t)] \\
&= W_{i2}(t) n^{1/2}[\widehat{\Lambda}_i^C(t+T_{i1}|T_{i1}) - \Lambda_i^C(t+T_{i1}|T_{i1})] + o_p(1) \\
&= W_{i2}(t) \left[\boldsymbol{k}'_{iC}(t+T_{i1};\boldsymbol{\alpha}) \boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1} n^{-1/2} \sum_{i=1}^{n} \int_0^{\tau} \{\boldsymbol{Z}_i^C - \overline{\boldsymbol{z}}_C(t;\boldsymbol{\alpha})\} dM_i^C(t) \right. \\
&\quad \left. + \exp\{\boldsymbol{\alpha}' \mathbf{Z}_i^C\} n^{-1/2} \sum_{i=1}^{n} \int_0^{t+T_{i1}} s_C^{(0)}(s;\boldsymbol{\alpha})^{-1} dM_i^C(s;\boldsymbol{\alpha}) \right] + o_p(1).
\end{aligned}
$$

B.3.4 $n^{1/2}[\widehat{A}_n(t;\widehat{W}_{i2}) - A(t)]$

We denote that

$$
A(t) = \int_0^t E[W_{i2}(s;t_1) \exp\{\boldsymbol{\beta}' \mathbf{Z}_i\}] d\Lambda_{02}(s;t_1).
$$

Then, the quantity of interest can be decomposed into two parts:

$$
\begin{aligned}
& n^{1/2}[\widehat{A}_n(t;\widehat{W}_{i2}) - A(t)] \\
&= n^{1/2}[\widehat{A}_n(t;\widehat{W}_{i2}) - \widehat{A}_n(t;W_{i2})] && \text{(B.3)} \\
&\quad + n^{1/2}[\widehat{A}_n(t;W_{i2}) - A(t)]. && \text{(B.4)}
\end{aligned}
$$

The first part, (B.3), can be written as,

$$
\begin{aligned}
&n^{1/2}[\widehat{A}_n(t;\widehat{W}_{i2}) - \widehat{A}_n(t;W_{i2})] \\
&=n^{-1}\sum_{i=1}^{n}\int_0^t n^{1/2}[\widehat{W}_{i2}(s;t_1) - W_{i2}(s;t_1)]N_{i2}(ds;t_1) \\
&=n^{-1}\sum_{i=1}^{n}\int_0^t W_{i2}(s)\left[\boldsymbol{k}'_{iC}(s+T_{i1};\boldsymbol{\alpha})\boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1}n^{-1/2}\sum_{j=1}^{n}\int_0^\tau\{\boldsymbol{Z}_j^C - \overline{\boldsymbol{z}}_C(u;\boldsymbol{\alpha})\}\right. \\
&\quad \left. dM_j^C(u;\boldsymbol{\alpha}) + \exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}n^{-1/2}\sum_{j=1}^{n}\int_0^{s+T_{i1}} s_C^{(0)}(u)^{-1}dM_j^C(u)\right]N_{i2}(ds;t_1) + o_p(1) \\
&=n^{-1/2}\sum_{j=1}^{n}\left\{\int_0^t n^{-1}\sum_{i=1}^{n}W_{i2}(s)\boldsymbol{k}'_{iC}(s+T_{i1};\boldsymbol{\alpha})\boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1}\int_0^\tau\{\boldsymbol{Z}_j^C - \overline{\boldsymbol{z}}_C(u;\boldsymbol{\alpha})\}\right. \\
&\quad dM_j^C(u)N_{i2}(ds;t_1) + \int_0^t n^{-1}\sum_{i=1}^{n}W_{i2}(s)\exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}\int_0^{s+T_{i1}} s_C^{(0)}(u;\boldsymbol{\alpha})^{-1} \\
&\quad \left. dM_j^C(u;\boldsymbol{\alpha})N_{i2}(ds;t_1)\right\} + o_p(1).
\end{aligned}
$$

The second part, (B.4), can be expressed as,

$$
\begin{aligned}
&n^{1/2}[\widehat{A}_n(t;W_{i2}) - A(t)] \\
&=n^{-1/2}\sum_{i=1}^{n}\int_0^t W_{i2}(s;t_1)N_{i2}(ds;t_1) - n^{1/2}\int_0^t E[W_{i2}(s;t_1)\exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}]d\Lambda_{02}(s;t_1) \\
&=n^{-1/2}\sum_{i=1}^{n}\int_0^t W_{i2}(s;t_1)N_{i2}(ds;t_1) \\
&\quad - n^{-1/2}\sum_{i=1}^{n}\int_0^t W_{i2}(s;t_1)\exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}d\Lambda_{02}(s;t_1) + o_p(1) \\
&=n^{-1/2}\sum_{i=1}^{n}\int_0^t W_{i2}(s;t_1)\left[N_{i2}(ds;t_1) - \exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}d\Lambda_{02}(s;t_1)\right] + o_p(1) \\
&=n^{-1/2}\sum_{i=1}^{n}\int_0^t W_{i2}(s;t_1)M_{i2}(ds;t_1) + o_p(1).
\end{aligned}
$$

Combining the results, we have

$$
\begin{aligned}
&n^{1/2}[\widehat{A}_n(t;\widehat{W}_{i2}) - A(t)] \\
&=n^{-1/2}\sum_{j=1}^{n}\left\{\int_0^t n^{-1}\sum_{i=1}^{n} W_{i2}(s)\boldsymbol{k}'_{iC}(s+T_{i1};\boldsymbol{\alpha})\boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1}\int_0^\tau \{\boldsymbol{Z}_j^C - \overline{\boldsymbol{z}}_C(u)\}dM_j^C(u)\right. \\
&\quad \left. N_{i2}(ds;t_1) + \int_0^t n^{-1}\sum_{i=1}^{n} W_{i2}(s)\exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}\int_0^{s+T_{i1}} s_C^{(0)}(u)^{-1}dM_j^C(u)N_{i2}(ds;t_1)\right\} \\
&\quad + n^{-1/2}\sum_{i=1}^{n}\int_0^t W_{i2}(s;t_1)M_{i2}(ds;t_1) + o_p(1) \\
&\equiv n^{-1/2}\sum_{j=1}^{n}\varphi_{j1}(t) + o_p(1).
\end{aligned}
$$

B.3.5 $n^{1/2}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$

By the asymptotic results of the Cox model, we have,

$$
n^{1/2}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \boldsymbol{\Sigma}_1(\boldsymbol{\beta})^{-1}n^{-1/2}\sum_{i=1}^{n}\int_0^\tau \{\boldsymbol{Z}_i - \overline{\boldsymbol{z}}_1(t;\boldsymbol{\beta})\}dM_{i1}(t;\boldsymbol{\beta}) + o_p(1).
$$

B.3.6 $n^{1/2}[\widehat{\Lambda}_{i1}(t;\widehat{\boldsymbol{\beta}}) - \Lambda_{i1}(t)]$

Similarly with Section B.3.2, we derive that

$$
\begin{aligned}
n^{1/2}[\widehat{\Lambda}_{i1}(t;\widehat{\boldsymbol{\beta}}) - \Lambda_{i1}(t)] =& \boldsymbol{k}'_{i1}(t;\boldsymbol{\beta})\boldsymbol{\Sigma}_1(\boldsymbol{\beta})^{-1}n^{-1/2}\sum_{i=1}^{n}\int_0^\tau \{\boldsymbol{Z}_i - \overline{\boldsymbol{z}}_1(t;\boldsymbol{\beta})\}dM_{i1}(t;\boldsymbol{\beta}) \\
&+ \exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}n^{-1/2}\sum_{i=1}^{n}\int_0^t s_1^{(0)}(s;\boldsymbol{\beta})^{-1}dM_{i1}(s;\boldsymbol{\beta}) + o_p(1),
\end{aligned}
$$

where

$$
\boldsymbol{k}_{i1}(t;\boldsymbol{\beta}) = \exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}\int_0^t \{\mathbf{Z}_i - \overline{\mathbf{z}}_1(t;\boldsymbol{\beta})\}\, d\Lambda_{01}(t).
$$

B.3.7 $n^{1/2}[\widehat{B}_n(t;\widehat{W}_{i2},\widehat{\Lambda}_{i1}) - B(t;W_{i2},\Lambda_{i1})]$

We denote

$$
B(t) = \int_0^t E[W_{i2}(s;t_1)\exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}]d\Lambda_{01}(s).
$$

Thus, it can be decomposed that

$$n^{1/2}[\widehat{B}_n(t;\widehat{W}_{i2},\widehat{\Lambda}_{i1}) - B(t;W_{i2},\Lambda_{i1})]$$

$$= n^{1/2}[\widehat{B}_n(t;\widehat{W}_{i2},\widehat{\Lambda}_{i1}) - \widehat{B}_n(t;W_{i2},\widehat{\Lambda}_{i1})] \tag{B.5}$$

$$+ n^{1/2}[\widehat{B}_n(t;W_{i2},\widehat{\Lambda}_{i1}) - B(t;W_{i2},\Lambda_{i1})]. \tag{B.6}$$

The first part, (B.5), can be written as

$$\begin{aligned}
&n^{1/2}[\widehat{B}_n(t;\widehat{W}_{i2},\widehat{\Lambda}_{i1}) - \widehat{B}_n(t;W_{i2},\widehat{\Lambda}_{i1})] \\
&= n^{-1}\sum_{i=1}^{n}\int_0^t n^{1/2}[\widehat{W}_{i2}(s;t_1) - W_{i2}(s;t_1)]\widehat{\Lambda}_{i1}(ds) \\
&= n^{-1}\sum_{i=1}^{n}\int_0^t W_{i2}(s)\left[\boldsymbol{k}'_{iC}(s+T_{i1})\boldsymbol{\Sigma}_C(\boldsymbol{\alpha})^{-1}n^{-1/2}\sum_{j=1}^{n}\int_0^\tau\{\boldsymbol{Z}_j^C - \overline{\boldsymbol{z}}_C(u;\boldsymbol{\alpha})\}dM_j^C(u)\right. \\
&\quad \left. + \exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}n^{-1/2}\sum_{j=1}^{n}\int_0^{s+T_{i1}} s_C^{(0)}(u;\boldsymbol{\alpha})^{-1}dM_j^C(u;\boldsymbol{\alpha})\right]\Lambda_{i1}(ds) + o_p(1) \\
&= n^{-1/2}\sum_{j=1}^{n}\left\{\int_0^t n^{-1}\sum_{i=1}^{n} W_{i2}(s)\boldsymbol{k}'_{iC}(s+T_{i1})\boldsymbol{\Sigma}_C^{-1}\int_0^\tau\{\boldsymbol{Z}_j^C - \overline{\boldsymbol{z}}_C(u)\}dM_j^C(u)\Lambda_{i1}(ds)\right. \\
&\quad \left. + \int_0^t n^{-1}\sum_{i=1}^{n} W_{i2}(s)\exp\{\boldsymbol{\alpha}'\mathbf{Z}_i^C\}\int_0^{s+T_{i1}} s_C^{(0)}(u;\boldsymbol{\alpha})^{-1}dM_j^C(u;\boldsymbol{\alpha})\Lambda_{i1}(ds)\right\} + o_p(1) \\
&\equiv n^{-1/2}\sum_{j=1}^{n} q_{j1}(t) + o_p(1).
\end{aligned}$$

The second part, (B.6), is

$$
\begin{aligned}
&n^{1/2}[\widehat{B}_n(t;W_{i2},\widehat{\Lambda}_{i1})-B(t;W_{i2},\Lambda_{i1})]\\
&=\frac{1}{n}\sum_{i=1}^{n}\int_0^t W_{i2}(s;t_1)n^{1/2}[\widehat{\Lambda}_{i1}(ds)-\Lambda_{i1}(ds)]\\
&=\frac{1}{n}\sum_{i=1}^{n}\int_0^t W_{i2}(s;t_1)\left[e^{\boldsymbol{\beta}'\mathbf{Z}_i}[\mathbf{Z}_i-\overline{\boldsymbol{z}}_1(s)]'\Lambda_{01}(ds)\boldsymbol{\Sigma}_1(\boldsymbol{\beta})^{-1}n^{-\frac{1}{2}}\sum_{j=1}^{n}\int_0^\tau\{\boldsymbol{Z}_j-\overline{\boldsymbol{z}}_1(u)\}\right.\\
&\quad\left.dM_{j1}(u)+\exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}n^{-\frac{1}{2}}\sum_{j=1}^{n}s_1^{(0)}(s;\boldsymbol{\beta})^{-1}dM_{j1}(s;\boldsymbol{\beta})\right]+o_p(1)\\
&=n^{-\frac{1}{2}}\sum_{j=1}^{n}\left\{\int_0^t\frac{1}{n}\sum_{i=1}^{n}W_{i2}(s;t_1)e^{\boldsymbol{\beta}'\mathbf{Z}_i}[\mathbf{Z}_i-\overline{\boldsymbol{z}}_1(s;\boldsymbol{\beta})]'\boldsymbol{\Sigma}_1(\boldsymbol{\beta})^{-1}\int_0^\tau\{\boldsymbol{Z}_j-\overline{\boldsymbol{z}}_1(u;\boldsymbol{\beta})\}\right.\\
&\quad\left.dM_{j1}(u)\Lambda_{01}(ds)+\int_0^t\frac{1}{n}\sum_{i=1}^{n}W_{i2}(s;t_1)\exp\{\boldsymbol{\beta}'\mathbf{Z}_i\}s_1^{(0)}(s)^{-1}dM_{j1}(s)\right\}+o_p(1)\\
&\equiv n^{-1/2}\sum_{i=1}^{n}q_{j2}(t)+o_p(1).
\end{aligned}
$$

Combining the two parts, we have

$$
n^{1/2}[\widehat{B}_n(t;\widehat{W}_{i2},\widehat{\Lambda}_{i1})-B(t;W_{i2},\Lambda_{i1})]=n^{-1/2}\sum_{i=1}^{n}\varphi_{i2}(t)+o_p(1),
$$

where $\varphi_{i2}(t)=q_{i1}(t)+q_{i2}(t)$.

B.3.8 $n^{1/2}[\exp\{\widehat{\theta}\}-\exp\{\theta\}]$

By a Taylor series expansion, along with the previous proved results, we have

$$
\begin{aligned}
&n^{1/2}[\exp\{\widehat{\theta}\}-\exp\{\theta\}]\\
&=n^{1/2}\left[\frac{\widehat{A}_n(t;\widehat{W}_{i2})}{\widehat{B}_n(t;\widehat{W}_{i2},\widehat{\Lambda}_{i1})}-\frac{A(t)}{B(t)}\right]\\
&=\frac{1}{B(t)}n^{1/2}[\widehat{A}_n(t;\widehat{W}_{i2})-A(t)]-\frac{A(t)}{B(t)^2}n^{1/2}[\widehat{B}_n(t;\widehat{W}_{i2},\widehat{\Lambda}_{i1})-B(t)]+o_p(1)\\
&=\frac{1}{B(t)}n^{-1/2}\sum_{i=1}^{n}\varphi_{i1}(t)-\frac{A(t)}{B(t)^2}n^{-1/2}\sum_{i=1}^{n}\varphi_{i2}(t)+o_p(1)\\
&=n^{-1/2}\sum_{i=1}^{n}\varphi_{i3}(t)+o_p(1),
\end{aligned}
$$

where

$$\varphi_{i3}(t) = \frac{1}{B(t)}\varphi_{i1}(t) - \frac{A(t)}{B(t)^2}\varphi_{i2}(t).$$

B.3.9 $n^{1/2}(\widehat{\theta} - \theta)$

We can write

$$n^{1/2}(\widehat{\theta} - \theta) = n^{1/2}[\log(e^{\widehat{\theta}}) - \log(e^{\theta})].$$

By a Taylor series expansion at e^{θ} and the results in Section B.3.8, we have

$$\begin{aligned} n^{1/2}(\widehat{\theta} - \theta) =& n^{1/2}[\log(e^{\widehat{\theta}}) - \log(e^{\theta})] \\ =& e^{-\theta} n^{1/2}(e^{\widehat{\theta}} - e^{\theta}) + o_p(1) \\ =& n^{-1/2} \sum_{i=1}^{n} e^{-\theta}\varphi_{i3}(t) + o_p(1) \\ \equiv& n^{-1/2} \sum_{i=1}^{n} \phi_i(t) + o_p(1), \end{aligned}$$

where $\phi_i(t) = e^{-\theta}\varphi_{i3}(t)$.

Thus, we have proved that under the regularity conditions we considered, $n^{1/2}(\widehat{\theta} - \theta)$ has a linear representation of $n^{-1/2}\sum_{i=1}^{n}\phi_i(t)$ asymptotically, where $\phi_i(t)$ $(i = 1, \cdots, n)$ are independent and identically distributed mean-zero random variables, such that $E\{\phi_i(t)^2\} < \infty$. Thus, $n^{1/2}(\widehat{\theta} - \theta)$ is asymptotically normal with mean 0 and variance $E\{\phi_i(t)^2\}$. The consistent estimator of the asymptotic variance is $\sum_{i=1}^{n}\widehat{\phi}_i(t)^2/n^2$, where

$$\widehat{\phi}_i(t) = e^{-\widehat{\theta}}[\frac{1}{\widehat{B}(t)}\widehat{\varphi}_{i1}(t) - \frac{\widehat{A}(t)}{\widehat{B}(t)^2}\widehat{\varphi}_{i2}(t)],$$

with

$$\widehat{A}(t) = \int_0^t n^{-1} \sum_{i=1}^{n} \left[\widehat{W}_{i2}(s;t_1) \exp\{\widehat{\boldsymbol{\beta}}' \mathbf{Z}_i\} \right] d\widehat{\Lambda}_{02}(s;t_1)$$

$$\widehat{B}(t) = \int_0^t n^{-1} \sum_{i=1}^{n} \left[\widehat{W}_{i2}(s;t_1) \exp\{\widehat{\boldsymbol{\beta}}' \mathbf{Z}_i\} \right] d\widehat{\Lambda}_{01}(s)$$

$$\begin{aligned}\widehat{\varphi}_{i1}(t) = &\int_0^t n^{-1} \sum_{j=1}^{n} \widehat{W}_{j2}(s) \Big\{ \widehat{\boldsymbol{k}}'_{jC}(s+T_{j1};\widehat{\boldsymbol{\alpha}}) \widehat{\boldsymbol{\Sigma}}_C(\widehat{\boldsymbol{\alpha}})^{-1} \widehat{\mathbf{U}}_i^C(\widehat{\boldsymbol{\alpha}}) + \exp\{\widehat{\boldsymbol{\alpha}}' \mathbf{Z}_j^C\} \\ &\int_0^{s+T_{j1}} S_C^{(0)}(u)^{-1} d\widehat{M}_i^C(u) \Big\} N_{j2}(ds;t_1) + \int_0^t \widehat{W}_{i2}(s;t_1) \widehat{M}_{i2}(ds;t_1)\end{aligned}$$

$$\begin{aligned}\widehat{\varphi}_{i2}(t) = &\int_0^t n^{-1} \sum_{j=1}^{n} \widehat{W}_{j2}(s) \Big\{ \widehat{\boldsymbol{k}}'_{jC}(s+T_{j1};\widehat{\boldsymbol{\alpha}}) \widehat{\boldsymbol{\Sigma}}_C(\widehat{\boldsymbol{\alpha}})^{-1} \widehat{\mathbf{U}}_i^C(\widehat{\boldsymbol{\alpha}}) \\ &+ \exp\{\widehat{\boldsymbol{\alpha}}' \mathbf{Z}_j^C\} \int_0^{s+T_{j1}} S_C^{(0)}(u;\widehat{\boldsymbol{\alpha}})^{-1} d\widehat{M}_i^C(u;\widehat{\boldsymbol{\alpha}}) \Big\} \widehat{\Lambda}_{j1}(ds) \\ &+ \int_0^t \frac{1}{n} \sum_{j=1}^{n} \widehat{W}_{j2}(s;t_1) \exp\{\widehat{\boldsymbol{\beta}}' \mathbf{Z}_j\} \Big\{ (\mathbf{Z}_j - \overline{\boldsymbol{z}}_1(s;\widehat{\boldsymbol{\beta}}))' \widehat{\boldsymbol{\Sigma}}_1(\widehat{\boldsymbol{\beta}})^{-1} \widehat{\mathbf{U}}_{i1}(\widehat{\boldsymbol{\beta}}) \widehat{\Lambda}_{01}(ds) \\ &+ S_1^{(0)}(s;\widehat{\boldsymbol{\beta}})^{-1} d\widehat{M}_{i1}(s;\widehat{\boldsymbol{\beta}}) \Big\},\end{aligned}$$

and

$$\widehat{\boldsymbol{k}}_{jC}(t;\boldsymbol{\alpha}) = \exp\{\widehat{\boldsymbol{\alpha}}' \mathbf{Z}_j^C\} \int_0^t \left\{ \mathbf{Z}_j^C - \overline{\mathbf{z}}_C(t;\widehat{\boldsymbol{\alpha}}) \right\} d\widehat{\Lambda}_0^C(t)$$

$$\widehat{\mathbf{U}}_i^C(\widehat{\boldsymbol{\alpha}}) = \int_0^\tau \{ \boldsymbol{Z}_i^C - \overline{\boldsymbol{z}}_C(t;\widehat{\boldsymbol{\alpha}}) \} d\widehat{M}_i^C(t;\widehat{\boldsymbol{\alpha}})$$

$$\widehat{\mathbf{U}}_{i1}(\widehat{\boldsymbol{\beta}}) = \int_0^\tau \{ \boldsymbol{Z}_i - \overline{\boldsymbol{z}}_1(t;\widehat{\boldsymbol{\beta}}) \} d\widehat{M}_{i1}(t;\widehat{\boldsymbol{\beta}}).$$

BIBLIOGRAPHY

BIBLIOGRAPHY

Ahmed, K., Ahmad, N., Khan, M. S., Koffman, G., Calder, F., Taylor, J. et al. *Influence of number of retransplants on renal graft outcome.* Transplant Proceedings 40, 1349–1352, 2008.

Andersen, P.K. and Gill, R.D. *Cox's regression model for counting processes: A large sample study.* The Annals of Statistics 10, 1100–1120, 1982.

Andrei, A.-C. and Murray, S. *Estimating the quality-of-life-adjusted gap time distribution of successive events subject to censoring.* Biometrika 93, 343–355, 2006.

Barba Abad, J., Robles García, J. E., Saiz Sansi, A., Tolosa Eizaguirre, E., Romero Vargas, L., Algarra Navarro, R. et al. *Impact of renal retransplantation on graft and recipient survival.* Arch Esp Urol 64, 363–370, 2011.

Breslow, N. E. *Contribution to the discussion of paper by D.R. Cox.* Journal of the Royal Statistical Society (Series B) 34, 216–217, 1972.

Chen, P.-Y. and Tsiatis, A. A. *Causal inference on the difference of the restricted mean lifetime between two groups.* Biometrics 57, 1030–1038, 2001.

Chen, Y. Q., Wang, M.-C., and Huang, Y. *Semiparametirc regression analysis on longitudinal pattern of recurrent gap times.* Biostatistics 5, 277–290, 2004.

Clement, D. Y. and Strawderman, R. L. *Conditional GEE for recurrent event gap times.* Biostatistics 10, 451–467, 2009.

Cox, D. R. *Regression models and life-tables.* Journal of the Royal Statistical Society (Series B) 34, 187–220, 1972.

Du, P., Jiang, Y., and Wang, Y. *Smoothing spline ANOVA frailty model for recurrent event data.* Biometrics 67, 1330–1339, 2011.

Duran, F.G., Cercadillo, R.A., Santos, L., de Diego, A., Ferreiroa, J., Valdecantos, E. *et al. Late orthotopic liver retransplantation: indications and survival.* Transplant Proceedings 30, 1876–1877, 1998.

Ghobrial, R.M. *Retransplantation for recurrent hepatitis C.* Liver Transplantation 8, S38–43, 2002.

Gong, Q. and Schaubel, D. E. *Partly conditional estimation of the effect of a time-dependent factor in the presence of dependent censoring.* Biometrics 69, 338–347, 2013.

Gruber, S. A., Brown, K. L., El-Amm, J. M., Singh, A., Mehta, K., Morawski, K. et al. *Equivalent outcomes with primary and retransplantation in African-American deceased-donor renal allograft recipients.* Surgery 146, 646–652; discussion 652–653, 2009.

Huang, X. and Liu, L. *A joint frailty model for survival time and gap times between recurrent events.* Biometrics 63, 389–397, 2007.

Huang, Y. *Two-sample multistage accelerated sojourn times model.* Journal of the American Statistical Association 95, 619–627, 2000.

Huang, Y. *Calibration regression of censored lifetime medical cost.* Journal of the American Statistical Association 97, 318–327, 2002.

Huang, Y. and Chen, Y.Q. *Marginal regression of gaps between recurrent events.* Lifetime Data Analysis 9, 293–303, 2003.

Irwin, J. O. *The standard error of an estimate of expectational life.* Journal of Hygiene 47, 188–189, 1949.

Jirka, J., Reneltová, I., Rossmann, P., Skibová, J., Macurová, H., Chadimová M. et al. *Repeat kidney transplantation.* Cas Lek Cesk 137, 686–689, 1998.

Kaplan, E. L. and Meier, P. *Nonparametric estimation from incomplete observations.* Journal of the American Statistical Association 53, 457–481, 1958.

Karrison, T. *Restricted mean life with adjustment for covariates.* Journal of the American Statistical Association 18, 151–167, 1987.

Kim, W.R., Smith, J.M., Skeans, M.A., Schladt, D.P., Harper, A.M., Wainright, J.L., et al. *OPTN/SRTR 2012 annual data report: Liver.* American Journal of Transplantation 14, 69–96, 2014.

Lin, D. Y., Sun, W., and Ying, Z. *Nonparametric estimation of the gap time distributions for serial events with censored data.* Biometrika 86, 59–70, 1999.

Little, R.J.A. and Rubin, D.B. *Statistical Analysis with Missing Data.* 2nd ed. New York: John Wiley, 2002.

United Network for Organ Sharing (UNOS). Data. UNOS Web site. Updated August 30, 2013. Available at http://www.unos.org/donation/index.php?topic=data. Accessed Sep 5, 2013.

Klein, J. P. and Moeschberger, M. L. *Survival analysis: techniques for censored and truncated data.* 2nd ed. New York: Springer, 2003.

Markmann, J.F., Markowitz, J.S., Yersiz, H., Morrisey, M., Farmer, D.G., Farmer, D.A., et al. *Long-term survival after retransplantation of the liver.* Annals of Surgery 226, 408–418; discussion 418–420, 1997.

Peña, E. A., Strawderman, R. L., and Hollander, M. *Nonparametric estimation with recurrent event data.* Journal of the American Statistical Association 96, 1299–1315, 2001.

Pour-Reza-Gholi, F., Nafar, M., Saeedinia, A., Farrokhi, F., Firouzan, A., Simforoosh N. et al. *Kidney retransplantation in comparison with first kidney transplantation.* Transplant Proceedings 37, 2962–2964, 2005.

Prentice, R. L., Williams, B. J., and Peterson, A. V. *On the regression analysis of multivariate failure time data.* Biometrika 68, 373–379, 1981.

Rao, P.S. and Ojo, A. *Organ retransplantation in the United States: Trends and implications.* Clinical Transplantation, 57–67, 2008.

Robins, J. M. and Wang, N. *Inference for imputation estimators.* Biometrika 87, 113–124, 2000.

Rosen, H.R. and Martin, P. *Hepatitis C infection in patients undergoing liver retransplantation.* Transplantation 66, 1612–1616, 1998.

Schaubel, D. E. and Cai, J. *Non-parametric estimation of gap time survival functions for ordered multivariate failure time data.* Statistics in Medicine 23, 1885–1900, 2004a.

Schaubel, D. E. and Cai, J. *Regression methods for gap time hazard functions for sequentially ordered multivariate failure time data.* Biometrika 91, 291–303, 2004b.

Strawderman, R. L. *The accelerated gap times model.* Biometrika 92, 647–666, 2005.

Strawderman, R. L. *A regression model for dependent gap times.* The International Journal of Biostatistics 2, ISSN (Online) 1557-4679, 2006.

Tejani, A. and Sullivan, E. K. *Factors that impact on the outcome of second renal transplants in children.* Transplantation 62, 606–611, 1996.

van der Laan, M. J., Hubbard, A. E., and Robins, J. M. *Locally efficient estimation of a multivariate survival function in longitudinal studies.* Journal of the American Statistical Association 97, 494–507, 2002.

Visser, M. *Nonparametric estimation of the bivariate survival function with an application to vertically transmitted AIDS.* Biometrika 83, 507–518, 1996.

Wang, M.-C. *Gap time bias in incident and prevalent cohorts.* Statistica Sinica 9, 999–1010, 1999.

Wang, M.-C. and Chang, S.-H. *Nonparametric estimation of a recurrent survival function.* Journal of the American Statistical Association 94, 146–153, 1999.

Wang, N. and Robins, J. M. *Large-sample theory for parametric multiple imputation estimators* Biometrika 85, 935–948, 1998.

Wang, W. and Wells, M. T. *Nonparametric estimation of successive duration times under dependent censoring.* Biometrika 85, 561–572, 1998.

Watt, K.D., Lyden, E.R., and McCashland, T.M. *Poor survival after liver retransplantation: Is hepatitis C to blame?* Liver Transplantation 9, 1019–1024, 2003.

Zucker, D. M. *Restricted mean life with covariates: Modification and extension of a useful survival analysis method.* Journal of the American Statistical Association 93, 702–709, 1998.

Zhang, M. and Schaubel, D. E. *Estimating differences in restricted mean lifetime using observational data subject to dependent censoring.* Biometrics 67, 740–749, 2011.

Zhang, M. and Schaubel, D. E. *Double-robust semiparametric estimator for differences in restricted mean lifetimes in observational studies.* Biometrics 68, 999–1009, 2012.